FLUORESCENCE LABELING AND COMPUTATIONAL ANALYSIS OF THE STRUT OF MYOSIN'S 50 kDa CLEFT

Ravi Kumar Gawalapu, M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2007

APPROVED:

Douglas D. Root, Major Professor
Robert M. Pirtle, Committee Member
Kent D. Chapman, Committee Member
Pamela A. Padilla, Committee Member
Thomas R. Cundari, Committee Member
Samuel F. Atkinson , Chair of the Department of Biological Sciences
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

Gawalapu, Ravi Kumar. Fluorescence labeling and computational analysis of the strut of myosin's 50 kDa cleft. Doctor of Philosophy (Biochemistry), August 2007, 148 pp., 3 tables, 60 illustrations, 120 references, 120 titles.

In order to understand the structural changes in myosin S1, fluorescence polarization and computational dynamics simulations were used. Dynamics simulations on the S1 motor domain indicated that significant flexibility was present throughout the molecular model. The constrained opening versus closing of the 50 kDa cleft appeared to induce opposite directions of movement in the lever arm. A sequence called the "strut" which traverses the 50 kDa cleft and may play an important role in positioning the actomyosin binding interface during actin binding is thought to be intimately linked to distant structural changes in the myosin's nucleotide cleft and neck regions. To study the dynamics of the strut region, a method of fluorescent labeling of the strut was discovered using the dye CY3. CY3 served as a hydrophobic tag for purification by hydrophobic interaction chromatography which enabled the separation of labeled and unlabeled species of S1 including a fraction labeled specifically at the strut sequence. The high specificity of labeling was verified by proteolytic digestions, gel electrophoresis, and mass spectroscopy. Analysis of the labeled S1 by collisional quenching, fluorescence polarization, and actin-activated ATPase activity were consistent with predictions from structural models of the probe's location. Although the fluorescent intensity of the CY3 was insensitive to actin binding, its fluorescence polarization was notably affected. Intriguingly, the mobility of the probe increases upon S1 binding to actin suggesting that the CY3 becomes displaced from interactions with the surface of S1 and is consistent with a structural change in the strut due to cleft motions. Labeling the strut reduced the affinity of S1 for actin but did not prevent actin-activated ATPase activity which makes it a potentially useful probe of the actomyosin interface. The different

conformations of myosin S1 indicated that the strut is not as flexible as several other key regions of myosin as determined by the application of force constraints to elastic portions of the myosin structure.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr Douglas D Root, for his outstanding guidance, insightful directions, enthusiasm, caring, encouragement and patience, providing me with an excellent atmosphere for doing research.

I would like to express my appreciation and gratitude to the members of my committee: Drs. Kent Chapman, Robert Pirtle, Pamela Padilla, Tom Cundari.

This dissertation could not have been written without the support and friendship found at University of North Texas and elsewhere. The love of family and friends provided my inspiration and was my driving force. It has been a long journey and completing this work is definitely a high point in my academic career. I could not have come this far without the assistance of many individuals and I want to express my deepest appreciation to them.

I would like to extend my thanks to former and present colleagues in the lab: Deepika, Zhilling, Dipesh, Pilar, Muthu and also I would like to thank Prem, Neal, Kim, Laurent, Catalina for their help and for enjoyable scientific discussion.

Finally, I would like to thank my parents, wife and my brothers and sister who have been a constant source of support and infinite love.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

ADP	adenosine 5’diphosphate
ATP	adenosine 5’triphosphate
DTT	dithiothreitol
EDTA	Ethylenediamine tetraacetic acid aminoethyl ether
ELC	essential light chain
HMM	heavy meromyosin
LC1	skeletal light chain 1
LC2	skeletal light chain 3
PMSF	phenyl methyl sufonyl fluoride
RLC	regulatory light chain
S1	myosin subfragment 1
RMSD	root mean square deviation

CHAPTER 1

INTRODUCTION

The Question

An understanding of the mechanism by which muscle generates force has been the subject of considerable study for decades. Key to our understanding of this mechanism is the conformational changes occurring in the muscle head as it interacts with the thin filament. This investigation was undertaken specifically to yield new insights into the responses communicated between distant sites in a skeletal muscle myosin II. In this dissertation, two techniques, fluorescence polarization and molecular dynamics simulations were used to measure conformational changes and to gain a detailed understanding of the structural changes in skeletal muscle myosin II. Herein are a description of the apparatus, the basis of the technique and results obtained from experiments performed on skeletal muscle myosin II. In order to put this work into context, literature regarding the biological and technical aspects of this research is reviewed in the following section.

Motor Proteins

Motor proteins are molecular machines that convert chemical energy from ATP hydrolysis into mechanical work, which powers cell motility. Molecular motors, in particular proteins of the myosin family, have attracted tremendous interest in the past decades because of its important role in cell motility and the interaction between myosin and actin filament. In 1993, the determination of the X-ray crystallographic structure of the S1 domain of myosin II was a big step for researches as it opened doors for understanding the motor proteins at the molecular level. Motor proteins are involved in a

large variety of varied cellular tasks. As a group, the motor proteins enable cells to move, divide, contract, secrete, cytokinesis, chromosome segregation, vesicle transport, endocytosis, organize the cytoplasm along with many others [1]. Among these apparently disparate events, the elements required for movement are quite similar. Thus information gleaned from the study of a specific motile event should shed light on the mechanism of many other processes.

The machinery of motility is commonly composed of two basic parts, tracks and motors. Tracks are filamentous protein polymers such as actin filaments and microtubules. There are three families of motor proteins that are involved in motion along cytoskeletal filaments: myosins, kinesins, and dyneins. Myosins move along actin filaments, whereas kinesins and dyneins translocate along microtubules. Today several other proteins are recognized as motors, which produce force and /or motion. Among these are F1ATPase [2], RNA polymerases [3], DNA helicases [4], the ribosomal complex including EF-Tu and EF-G [5] and several G-proteins [6]. Each of these proteins or protein complexes utilizes energy and undergoes a major structural change essential to its enzymatic function.

Examination of the crystal structures of members of the myosin and kinesin super families have shown that the motors have common structural motifs despite a lack of similarity in primary sequence. While the motor domain of myosin (~840 amino acids) is much larger than that of kinesin (~340 amino acids) the catalytic domains of these two enzymes are strikingly similar. They are both composed of a core of 8 β-sheets flanked by 7 α-helices, all of which overlap. The conserved structures encompass the machinery required for ATP hydrolysis. Among these structures are the phosphate binding loop (P

loop). The P-loop is found in these motors, many small G-proteins and several other proteins including adenylate kinase, transducin and recA. In addition to to the P-loop, myosins, kinesins and G-proteins have three other overlapping structures (commonly referred to as N1-4 or G1-4, where N1 is the P-loop, N2 is switch 1, N3 is switch 2 and N4 is a purine binding site) all of which play a role in nucleotide hydrolysis [7]. Although primary sequences of these proteins have limited similarity, key amino acids believed to play a direct role in hydrolysis, are conserved within the common structures. Based on these findings, it can be said that the mechanism of nucleotide hydrolysis has evolved from a common origin and has subsequently been specialized for specific nucleotides and conversion of the energy released into distinct signals [7].

While dynein also belongs to the superfamily of P-loop ATPases, structural similarities stop here. The motor unit is much larger (~1000 amino acids) including six 35-40 kDa modules, two of which have been shown to bind ATP and a microtubule binding domain at the end of a 100Å stalk.

Myosin Classes

The different motors belonging to the myosin superfamily are classified into 20 different myosin classes [8]. All myosins share the common feature that they consist of a globular motor domain (head domain), an α-helical neck domain that binds two light chains, and a long α-helical tail. Myosins diversified very early during eukaryotic evolution by a process of gene duplications and diversification from one progenitor molecule in the distant past, and it has recently been proposed that the eukaryotic cenancestor contained three different types of myosins [9,10]. Most eukaryotes require myosins, but a few groups, e.g, red algae and diplomonad protists live without them [11].

The different myosins have been distinguished on the basis of the sequence of amino acids in their ATP hydrolyzing motor domains. The different classes of myosin also differ in structure of their tail domains. Tail domains have various functions in different myosin classes, including dimerization and other protein-protein interactions. The best characterized myosins are the class II myosins. Myosin II, also termed as the conventional myosin is the founding member of the family, identified over 60 years ago. The most prominent among the class II myosins are the muscle myosins that form the thick filaments in a sarcomere of the muscle cell. Muscle contraction is facilitated by the rhythmic interaction between myosin heads and actin filaments (thin filaments) that result in relative movement of the thick and thin filaments. The myosin I family was discovered next and is named so because it has one catalytic domain in contrast to myosin II which has two [12]. The rest of the myosins are numbered in the order they were discovered .

The motor domain is considered to be the major part of the myosin molecule which exhibits a relatively conserved structure. Owing to the progress of various genome projects aiming at the deciphering of the species primary DNA sequences, the list of myosin members is continuously increasing. Functionally, the distinct myosins have been studied with respect to the following aspects:

1) *Step Direction*.

A key feature characteristic of a specific motor is its directionality, ie in which direction it moves along its track [13]. Until the finding of the walking direction of myosin VI, myosins were thought to walk along the same direction, from the pointed (-) end to the barbed (+) end of actin filament. However, the myosin VI molecule moves in the reverse direction and this might be responsible for its special role in transport *in vivo*.

2) *Step Size*

The average step length of distinct myosin species varies and is a unique property of each myosin member

3) *Power Stroke*.

The inner domains involved in force generation undergo conformation change with various displacements depending on different myosin species.

4) *Duty Ratio*.

The time that myosin molecule spends attached to the actin during the ATP catalytic cycle relative to the whole cycle time.

5) *Kinetics of ATP hydrolysis*

The rate constants of each step in the cross-bridge cycle, in particular the dissociation constants of the reaction products ADP/Pi, are of prime importance for the overall role of the motor.

6) *Processivity*

Processivity is defined as the average step number a given motor molecule remains bound to an actin filament per diffusion encounter . A high processivity of a motor protein enables itself to walk along its track for a relative long distance without dissociation from the track. Up to date, it is has been found that myosin V, VI, IX b and XI are processive molecules. Different motors can adopt different strategies on how to move along their filamentous tracks. Which strategy is used depends on the function of the motor. Motor proteins can move along their tracks either processively or non-processively. Myosin V motors for instance are involved in intracellular trafficking having vesicles as cargo to be transported along the actin cytoskeleton. It has been shown

that the myosin V dimer utilizes a mechanism in which one head of the myosin dimer remains attached to the track at all times [14]. Two scenarios are consistent with this, a hand-over-hand mechanism in which the two heads alternate in the lead or an inchworm mechanism in which one head always leads. Using high resolution fluorescence imaging techniques it was found that myosin V walks hand over hand [15]. Myosin II is a classical example of a non-processive motor that detaches from its track at each ATP turnover. This is adequate for myosin II because many myosin II monomers self assemble into filaments that slide relative to its track.

7) *Sliding Velocity*

In a well-defined *in vitro* system, a given kind of myosin molecule has a characteristic sliding velocity which is measured by motility assays. It is believed to be controlled by a combination of kinetic and mechanical properties. An extremely wide range (>4000 fold) from 0.015 μm/s (myosin IX) to70 μm/s (myosin XI) has been found.

8) *Regulation*

It refers both to the regulation of the internal conformational changes in the motor domain and the activation of the motor function (ATPase, power stroke etc.). Regulation may be achieved in several ways: by activation of myosin binding domain on the actin filament regulated by troponin/tropomyosin in the case of muscle derived myosin II or by light chain phosphorylation/dephosphorylation as in the case of myosin II smooth muscle cells, and/or by binding/unbinding as in myosin V. In all cases, Ca 2+ seems to play a central role as a signaling agent.

Unconventional myosins and their receptors are involved in diverse tasks such as organelle translocation and cytoskeletal reorganization (myosin I), cytokinesis

(nonmuscle myosin II), maintenance of cell integrity and structure and signaling (myosin III), vesicle transport and membrane trafficking (myosin V), particle transport and anchoring (myosin VI), cell adhesion, hearing, and maintenance of balance (myosin VII), signaling (myosin IX), filopod extension (myosin X), and gliding, motility (myosin XIV).

Several members of the myosin superfamily have been directly linked to diseases. Myosin VI and VII have been implicated in diseases that cause deafness and blindness [16] while myosin V may be essential to neurological function and has been linked to Griscelli syndrome [17]. Given that isoforms of myosin II are major components of smooth and skeletal muscle, it is not surprising that mutations in any of these isoforms can have dramatic effects. For example mutations of myosin II are linked to hypertrophic cardiomyopathy [18]. In all cases the patients are heterozygous for this mutation. Presumably, homozygous and not viable.

Size and Shape of Myosin II

Myosin II is a large asymmetric molecule, it has a long tail and two globular heads (Fig.1.1). The tail is about 1,600 Å long and 20 Å wide. Each head is about 165 Å long, 65 Å wide and 40 Å deep at its thickest part. The molecular weight of myosin is about 500,000 Da. In strong denaturing solutions, myosin dissociates into six polypeptide chains of two heavy chains and four light chains with each head containing two light chains. Each heavy chain weighs about 200,000 Da and the two regulatory light chains each weigh 20,000 Da, one essential light chain weighs around 15,000 Da and another 25,000 Da. The two heavy chains are wound around each other to form a double helical structure.

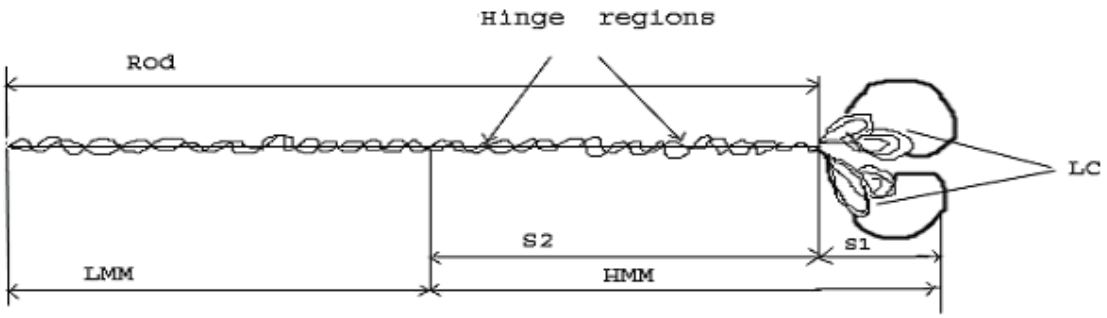

Fig1.1: Schematic representation of the molecule of muscle myosin II and its fragmentation by proteolytic enzymes (Redrawn from Rayment etal , 1993) : S1, myosin subfragment 1; S2, myosin subfragment 2; HMM, heavy meromyosin; LMM, light meromyosin; LC, light chains; Rod, rod part of the molecule. Hinge regions has low content of alpha-helix and so provides a high mobility of the myosin heads relative to the rod part and of the HMM region relative to the LMM region.

Myosin Structure

The motor activity of myosin head (S1) was demonstrated in 1987 when isolated myosin heads moved actin filaments in an *in vitro* motility assay [19]. It became clear that just the myosin head can undergo motile functions. Therefore, the interest of most investigators studying myosin structure and functions switched to studies on the myosin head. In 1993, the three-dimensional structure of the myosin head was determined by X-ray analysis of S1 crystals [20] which revealed the finer details of the myosin structure.

Myosins are typically composed of three functional domains: the head (motor domain, MD), neck (light chain domain, LCD) and tail domains. The S1 is elongated and consists of a seven stranded β sheet and a C-terminal α helical tail which carries the two calmodulin-like light chains (Fig.1.2). The MD is the actin binding site and nucleotide binding site. It is highly conserved across classes with the exception of a few variable surface loops and N-terminus. The LCD is an α-helical structure with light chains bound

to one or several repeating IQ-motifs. The IQ-motif is an 11 amino acid sequence with the conserved pattern IQxxxRGxxxr, and is similar to the calmodulin binding site of neuromodulin [21]. It was first recognized in myosins upon cloning of myosin V. Many known calmodulin binding proteins and myosins have at least one IQ-motif. Myosin light chains are calmodulin or calmodulin like proteins. These proteins contain EF-hand structures, which are Ca2+ and/or Mg2+ binding sites. They play a variety of roles including structural (stabilizing the α-helix of the heavy chain) and regulatory (e.g. smooth muscle myosin is inactive unless the light chain (RLC) is phosphorylated). The tail is the least conserved of myosin domains. Tails vary in structure and presumably function. Several classes of myosin have tails that form coiled coil structures resulting in a double-headed myosin molecule. In general, it is assumed that the tail plays an anchoring role connecting myosin to another structure such that this structure can be stabilized. In principle, the head could be an anchor for the tail. Other functional motifs, such as GTPase activation and plekstrin homology domains are found in several myosin tails.

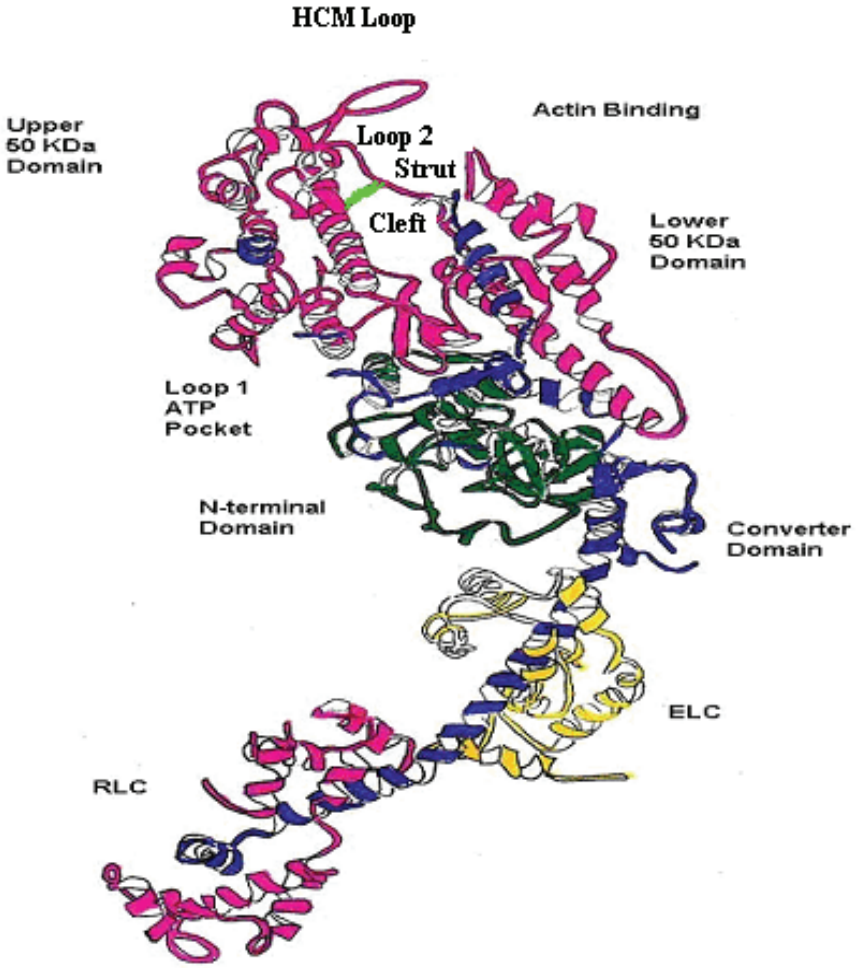

Fig.1.2: Three-dimensional structure of subfragment1 redrawn from (Rayment etal 1993). The picture shows the head of S1 consisting of a seven-stranded beta-sheet and a C-terminal tail containing the regulatory light chain (magenta) and the essential light chain (yellow). The proteolytic fragments of S1 are color coded as follows: 25K (N-terminal), green; 50K, red; and 20K (C-terminal) blue. The 50K fragment spans two domains: the 50K upper domain and the 50K lower domain or actin-binding domain. The strut is fluorescent green which connects the upper and lower 50 kDa domain. Part of the 50-kDa and 20-kDa fragments form the actin-binding site, whereas part of the 50-kDa and 25-kDa fragments of S1 form the ATP-binding site. The ATP-binding site is about 4 nm from the actin-binding site. There is a cleft in the upper part of the head that extends from under the ATP-binding site to the actin interface. Both portions, above and below the cleft are involved in the actin binding. When ATP binds to S1, the pocket most likely closes and the cleft widens disturbing the S1-actin binding, that is ATP dissociates S1 from actin. When ATP is hydrolyzed by S1 to ADP and P_i, actin recombines with S1. The accompanying structural changes are the narrowing of the cleft and opening of the ATP-binding pocket. These subtle changes are called conformational changes that play a key role in the mechanism of muscle contraction.

Partial digestion of S1 with trypsin, results in cleavage of the heavy chain at two regions, around residues 204-213 (loop1) and 636-642 (loop2) and upon denaturation fragments of a) 25K (N-terminal) b) 50K (middle) c) 20K (C-terminal) are obtained. The 50K fragment spans two domains, the 50K upper domain and the 50K lower domain (actin binding domain). The ATP-binding site lies near the 25K-50K boundary and contains a P-loop, which is also found in many ATPases and G-proteins. The 20K domain contains a broken helix with two reactive thiols (SH1 and SH2). The converter domain follows the SH1 helix and functions as a socket for the C-terminal helical tail. The C-terminal tail that is associated with the light chains is called the regulatory domain or neck and its main function is to amplify rotational movements by the converter domain [22-24].

The tails of the two heavy chains dimerize as an α-helical coiled-coil. The tails can further polymerize with other myosin molecules to make a bipolar thick filament. This structure is a major component of skeletal muscle. Polymerization is reversible and dependent on ionic strength. Thick filaments are not soluble in aqueous solution at physiological ionic strength, which severely hinders kinetic studies. Proteolytic fragments of myosin are soluble and retain enzymatic activity. Myosin is readily cleaved with α-chymotrypsin at two different sites resulting in single and double headed fragments. The single headed portion is referred to as subfragment-1(S1). It includes the MD, and part of the LCD with one or both light chains (depending on proteolysis conditions). It retains actin-activated ATPase activity and is sufficient to support in vitro motility [19]. The double-headed fragment is called heavy meromyosin (HMM). It retains enough of the coiled coil beyond the LCD to dimerize but not enough to allow further polymerization

when S1 is made, the remainder of the molecule is called the rod. When HMM is made only light meromyosin (LMM) remains. Both of these fragments are capable of polymerizing.

The motor activity of myosin is based on its interaction with actin, which is coupled with hydrolysis of ATP. The head of myosin, is where the sites responsible for the interaction of actin and nucleotide are located, including a large cleft with a strut sequence connecting the upper and lower halves [25-27]. The converter region in S1 changes structure upon binding actin and nucleotide, and the lever arm turns ~60 degrees relative to the motor domain [28]. The regulatory domain acts as a lever whose length determines the size of displacement of the myosin head along the actin filament. The length of the lever arm is found to be propotional to the speed of transport of actin filaments by immobilized S1 molecules in the in vitro motility assay [29].

The conformational changes in myosin catalysed ATP hydrolysis are shown by the simplified scheme of Bagshaw and Trentham [30]:

$$M + ATP \longleftrightarrow M^{*}\cdot ATP \longleftrightarrow M^{**}\cdot ADP\cdot P_i \longleftrightarrow M^{**}\cdot ADP + P_i \longleftrightarrow M + ADP + P_i$$

M, $M^{*}\cdot ATP$, $M^{**}\cdot ADP\cdot P_i$, and $M^{*}\cdot ADP$ represent myosin or S1–nucleotide complexes with distinct tryptophan fluorescence, which exist as intermediate states during the hydrolysis of ATP [31]. The conformational changes taking place during the transitions between the intermediate states of ATP hydrolysis are coupled with alterations in actin affinity and are essential for force generation. Therefore, the conformation of the myosin intermediate states during hydrolysis of ATP is necessary for understanding the mechanism of the motor activity.

Actin

In any eukaryotic cell, actin is the most abundant protein. Vertebrate skeletal actin has 375 residues and a molecular mass of 42.2 kDa. Actin, so named because it activates the ATPase activity of myosin, is a major component of the cytoskeleton within the cell. It is highly conserved in eukaryotic cells and binds to at least a hundred different proteins [32]. Actin filaments are dynamic polymers whose ATP-driven assembly in the cell cytoplasm drives shape changes, cell locomotion and chemotactic migration. Actin filaments also participate in muscle contraction.

The actin monomer, also called G-actin, is a globular protein separated into two lobes by a deep cleft and is ~50 Å by 55 Å by 33 Å. Actin polymerization is a self driven process where ATP hydrolysis is not necessary. The polymerization of actin monomers with bound ATP forms the actin filament (F-actin), which is a two state right handed helix, with a crossover every 13-14 actin monomers [33]. Besides appropriate pH adjusted by buffer, ions and ATP as well as the concentration of G-actin play critical roles in polymerization. Although there is no atomic model of the actin filament, a model of F-actin was made using parameters derived from the x-ray diffraction data of flow oriented F actin and the crystal structure of the actin monomer [34].

In vitro Motility Assays

The study of motor proteins was revolutionized by the development of in vitro motility assays in which the motility of purified motor proteins with purified cytoskeletal filaments is reconstituted in cell free conditions. In this assay, myosin is adhered to a nitrocellulose-coated microscope coverslip (Fig.1.3). The coverslip is one surface of a 20 μl experimental microchamber. Fluorescently labeled actin filaments in a solution

containing MgATP are perfused into the chamber and then the movement of single actin filaments over the myosin surface is visualized in an epifluorescence microscope equipped with an image intensified video camera. Actin filament velocities can be determined by tracking the leading edge of the fluorescent filaments from frame to frame using appropriate software.

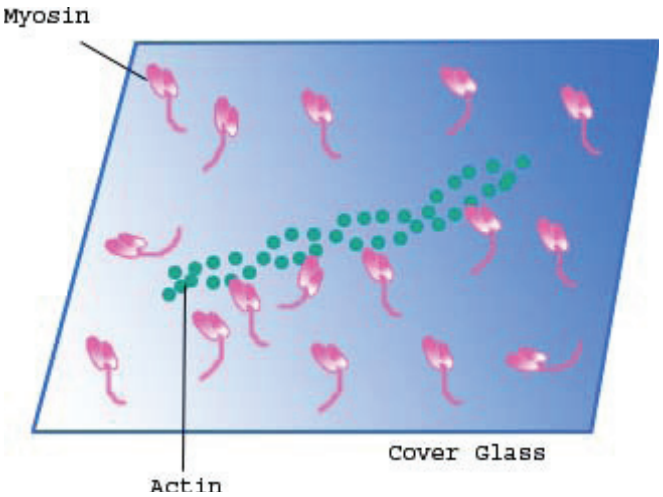

Fig.1.3: Schematic illustration of invitro motility assay

ATPase Activity of Myosins

To generate motility motor proteins convert chemical energy into mechanical energy by coupling the hydrolysis of ATP to conformational changes. Myosin II has served as a model system for understanding motility for decades. Unfortunately, an intact myosin II molecule is not well suited for kinetic studies of the MgATPase activity, particularly in presence of actin as it polymerizes into thick filaments at physiological and lower ionic strengths. These filaments do not mix well and thus contribute to large amount of light scattering. Furthermore, thick filaments interact heterogenously with actin filaments, complicating the kinetic analysis. Therefore, most of the data obtained on the kinetic cycle of myosins are from soluble proteolytic fragments, S1 or HMM. The

same is also true for unconventional myosins. Myosin S1 from vertebrate skeletal muscle hydrolyses MgATP slowly (kcat =0.05 s-1), but this activity is acclerated by two orders of magnitude by F-actin at low ionic strength. Although the steady-state ATPase activity of myosin is slow, the hydrolysis step itself is rapid (about 100 s-1) and leads to a longlived products complex (M.ADP.Pi) as the major steady-state intermediate.

Actomyosin Interface

To date it has not been possible to crystallize the actomyosin complex and produce a crystal structure of the actomyosin interface. An atomic resolution model of the actomyosin rigor complex has been constructed by fitting the electron density of the S1 crystal structure with a modeled F-actin structure [35].

The areas of contact between the myosin head and actin can be divided into four distinct regions. Primary actin contacts on S1 are amino acids 529-558 and amino acids 647-659. Secondary sites are an unstructured surface loop (aa 567-578) and the structured myopathy loop (aa 403-416) present on the S1 surface [36]. Myopathy loop is a structured surface loop containing an unusual clustering of heart disease-implicated mutations, implying a functional significance for the loop [37-39]. The myopathy loop is probably an actin binding site however, the role of its disease-implicated point mutations in motor functions remains unclear. The C-loop is the second structured surface loop to be implicated in cardiac myosin motor function.

The binding site appears to be centered on a region of hydrophobic interaction involving helices at the myosin and actin surface. This site is surrounded by both polar and ionic groups. On the myosin head, this region is flanked by surface loops on three sides that are well placed for ionic interactions with the surface of actin. It is also thought

that there may be hydrogen bonding between the polypeptide backbone on adjacent myosin and actin loops. Presently, we lack a detailed understanding of actomyosin structure and in particular how the elements making up the actomyosin interface contribute to mutual affinity, actin regulation of phosphate release, and the ability of myosin-bound nucleotide to modulate actomyosin affinity.

Analysis of the binding site is complicated because several of the surface loops on the S1 model are undefined [40,41,35] and that models for F-actin are still undergoing refinement. The binding affinity of myosin and actin varies with the chemical state of the bound nucleotide. As the biochemical state of the nucleotide changes both the affinity and location of the actin binding sites may change. The structural basis for the interaction of skeletal S1 with actin is provided, partly, by the 50 kDa lower domain, which was found to maintain substantial rigidity in the different nucleotide states.

Structural Studies of Myosin S1

It is well known that binding of actin and binding of nucleotides to myosin are antagonistic, which is the biochemical basis for the crossbridge cycle of muscle contraction. ATP binding to actomyosin causes actin dissociation, whereas actin binding to the myosin accelerates ADP and phosphate release. Hydrolysis of ATP provides the energy for changes in the conformation of the myosin head that occur, while it is attached to actin. The structural effect of actin causes decreased affinity of the products of ATP hydrolysis to myosin [42], which is the cause of the actin activation of myosin ATPase, changed proteolytic susceptibility of the connector loops in S1 [43] and decreased reactivity of Cys707 (SH1) [44]. Structural studies have indicated that communication between the actin and nucleotide binding sites involves the opening and closing of the

cleft between the upper and lower 50K domains of the myosin head. There is little change in secondary structure during ATPase, and studies on the orientation of labels bound to myosin SH1 reveal little change during contraction. These results suggest that myosin head is made up of independently folded regions that can move relative to each other. In addition, knowledge of the actin-induced conformational changes in S1 is essential for understanding the work of the myosin-actin motor system. However, it is more difficult to define accurately the structural changes in S1 caused by actin than those caused by nucleotides, because the S1–actin complex has not been crystallized. The regions of S1 which make contact with a single actin subunit are structures which encircle a cleft in the 50 kDa fragment that is thought to undergo opening and closing motions [45]. In view of the importance of the description of both the nucleotide and actin-induced conformational changes in S1, it was decided to investigate these changes in solution at the strut location of the myosin head.

Familial Hypertrophic Cardiomyopathy

Mutations in myosin S1 can lead to Hypertrophic cardiomyopathy. Hypertrophic cardiomyopathy is a group of heart disorders in which the walls of the heart ventricles become thickened. The stiff ventricle walls make it difficult for the heart to fill properly with blood. Hypertrophic cardiomyopathy is usually familial. Familial hypertrophic cardiomyopathy (FHC) occurs when an individual inherits one mutated gene from a parent [46]. The percentage of people who inherit this disease is between 0.2% and 0.1%. It occurs more often in men than in women. Patients with hypertrophic cardiomyopathy may not exhibit any signs or symptoms. When symptoms occur, they include fainting, heart palpitations, chest pain, shortness of breath, and heart failure. The disease is

clinically variable, ranging from benign to severe, and often results in sudden death in young athletes [47,48]. Symptoms may appear early in childhood or later in life. Hypertrophic cardiomyopathy is usually evident upon examination with a stethoscope because the heart sounds with this condition are easy to identify. Diagnostic tests such as an echocardiogram, electrocardiogram, or a chest x-ray are frequently used to confirm the condition.

The majority of patients with hypertrophic cardiomyopathy benefit from medications such as beta-blockers or calcium channel blockers to help reduce the heart's resistance to filling with blood between heartbeats. Surgery is only considered for those who don't respond to drug therapy. The thickened heart muscle is removed to improve the flow of blood from the heart.

The mechanism by which the FHC mutations in sarcomeric proteins cause hypertrophy is still unknown. Mutations in sarcomeric proteins are likely to compromise sarcomere function resulting in hypertrophic cardiomyopathy. Cardiac hypertrophy can be the result of a number of stimuli, including exercise, pressure and volume overload. It is possible that the mutant myosin protein decreases myocardial contractility of the myocyte and imposes stress on the cardiac myocytes, which induces a hypertrophic response.

FHC is also genetically heterogeneous, because it is caused by multiple mutations in at least seven different genes [49,50]. All mutations found to cause FHC are structural proteins of the sarcomere. Some of these mutations appear in clusters near functionally significant regions of the myosin motor domain (S1), such as the actin binding domain, the ATP catalytic domain, and the LC-binding domain [39]. A number of mutations have

also been described at the junction between S1 and the S2 [39]. The crystal structure of the skeletal myosin head shows that the Arg249Gln mutation is near the ATP binding site and the Arg403Gln is in the actin-binding region and Val606Met is located in the strut region. Conformational studies at these regions would help in understanding the mechanism leading to the cause of the inherited forms of hypertrophic cardiomyopathy and could advance drug development.

None of the known FHC myosin mutations occur in proposed critical regions of force production, but the proximity of some of the mutations suggests that the motor function of the mutant myosin may be impaired. The Arg249Gln mutation has a moderate phenotype with respect to incidence and age of sudden death [18]. The Arg403Gln mutation is associated with a high frequency of sudden death in young adults and heart failure of adults in their fourth decade [51,52,18]. The Val606Met mutation is typically associated with a benign form of the disease, and most patients survive beyond the sixth decade [18]. Val606 is in a helix-loop-helix motif just at the upstream of the junction of 50-20 kD subdomains of S1 which contain a part of actin binding interface of myosin molecule [20]. The effect of Val606Met mutation on myosin motor and enzymatic activities was mild. On the basis of the structural analysis of acto-S1 complex [39], there are no indications of the importance of this segment for actin-myosin interaction, therefore, the very mild modulation of myosin function by the Val606Met mutation would be reasonable.

Conformational Studies of Myosin S1

Protein conformational studies are of paramount importance to understand biomolecular interactions. Conformational changes in proteins occur because proteins are intrinsically flexible. These conformational changes may occur with only relatively small expenditures of energy. At the molecular structural level, conformational changes in single polypeptides are due to changes in main chain torsional angles and side chain orientations. The overall effect of such changes may be localized with reorientations of a few residues and small torsional changes in the regional main chain. Conformational changes in a protein is often caused by docking with other atoms and molecules. The biological significance for investigating these conformational changes is to understand the domain-domain interplay in multi-domain proteins like myosin where some of the domains bind to other proteins or where a disease related mutation is present. Analyzing structure, dynamics and interactions of fragments of these proteins will reveal the role they play in a healthy organism, the way they are involved in disease and therefore one can develop treatments for the diseases. The recent developments in experimental methods, especially cryo-electron microscopy, have revealed that large-molecular assemblies are also highly dynamic. While experiments can provide a tremendous source of information on these dynamical properties, computational methods must be employed to complement experimental observations to obtain new insights into the mechanism of these transformations that are presently inaccessible to experiments.

The usefulness of fluorescence in studying protein conformations is due to its sensitivity, kinetic resolution, and compatibility with both live cells and physiological assays. Conformational changes within the myosin head during steady-state ATP

hydrolysis have been detected by observing changes in the UV absorbance [53], intrinsic fluorescence [30,31,54], biochemical crosslinking [55,56], electron paramagnetic resonance spectroscopy [57-59], electron microscopy [60-62], NMR spectroscopy [63], fluorescent labeling [64], fluorescence energy transfer [65], small-angle X-ray scattering [66]. In addition, fluorescence polarization is highly sensitive to conformational changes in myosin [67-69]. The goal of these studies has been to correlate structural changes with specific kinetic intermediates and the force-generating step. A problem which limits the interpretation of some spectroscopic measurements is that the attached label can cause local modifications, so labeled sites in the motor domain producing fully functional myosin are few [70]. Due to the localized information provided by a spectroscopic probe, several different labels must be used to understand the complex molecular motions of myosin.

In order to visualize the conformational changes during muscle contraction, the myosin head has been crystallized in three different states. The prehydrolysis states are represented by phosphate and beryllium fluoride analogs (MgATPγS, MgAMP.PNP, MgADP.BeFx), posthydrolysis states by orthovanadate, aluminum fluoride analogs (MgADP.Vi, MgADP.AlF4) [71-73] and a third state is obtained when scallop muscle myosin is complexed with MgADP [41]. These states are differentiated depending on the substructure of the 50kDa cleft, the converter region, and the position of the lever arm. Based on the comparison of these three states it is concluded that myosin head consists of four subdomains which are (upper 50kDa subdomain, lower 50kDa subdomain, N terminal subdomain, and converter) (Fig. 1.4). Three joints connecting the subdomains play a key role in their movement and coordinate the conformational changes that occur

in the myosin head as it hydrolyzes ATP. These joints are: switch II connecting the upper and the lower 50 kDa subdomains at the bottom of the actin binding cleft separating these subdomains, the relay connects the lower 50 kDa subdomain to the converter, and two short α helices located near the converter region of myosin contains the reactive sulfhydryl groups, SH1 (Cys-707) and SH2 (Cys-697), SH1 helix connects the N terminal subdomain to converter, and the SH2 helix is located in the N terminal subdomain close to the ATPase site, which is between the N terminal subdomain and the upper 50 kDa subdomain.

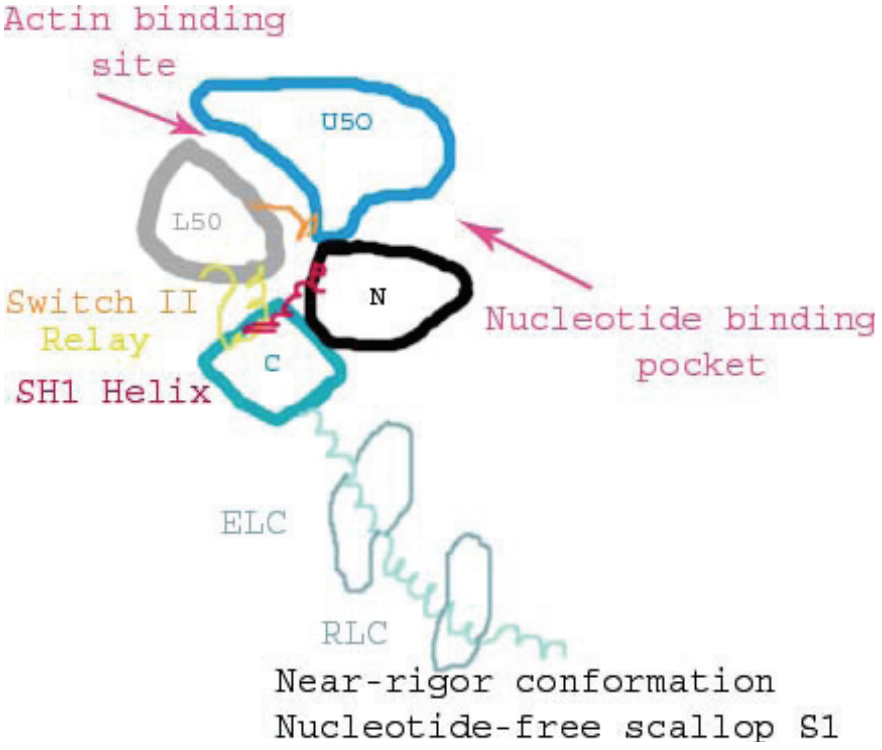

Fig.1.4 **:** Schematic drawing of the subdomains within the myosin head in the nucleotidefree scallop S1 structure (Redrawn from Houdusse et al, 2000). The four subdomains of the motordomain are labeled, as are three joints that articulate movements within the head: switch II (orange); the relay (yellow); and the SH1 helix (red): the N-terminal (black), upper 50-kDa subdomain (blue), lower 50-kDa subdomain (gray), and the converter (green) as well as for the three joints. ELC, essential light chain; RLC, regulatory light chain .

Research Objective

The primary motivation behind my studies is to explore, using computational and experimental approaches, functionally important rearrangements and to obtain new insights into the mechanism of these transformations that are presently inaccessible to experiments alone. Myosin is the molecule in muscle that does mechanical work. Every movement we make depends on myosin, including the beating of our hearts, breathing, talking, walking, reading (eye movements), writing (hand movements). To do work, myosin must convert chemical energy into mechanical work. Just how it does that is unknown, and is the fundamental question that is addressed.

The hypothesis for this dissertation is to test the mobility of the strut. The strut is thought to move during the energy transduction. This hypothesis is tested by using a novel fluorescent label to measure movement of the myosin strut. The movement of the strut is compared and analyzed with movements of an atomic model of myosin in different conformation states like opening of the 50 kDa cleft, closing of the 50 kDa cleft, opening of the nucleotide cleft, crosslinking of reactive cysteines (SH1 and SH2), opening the 50 kDa cleft with crosslinked cysteines, closing the 50 kDa cleft with crosslinked cysteines, opening the 50 kDa cleft with nucleotide cleft open, closing the 50 kDa cleft with nucleotide cleft open, opening the 50 kDa cleft with crosslinked cysteines and nucleotide cleft opened, closing the 50 kDa cleft with crosslinked cysteines and nucleotide cleft opened. This powerful combination of spectroscopy with modeling lead to a striking conclusion: that movements of myosin strut that might be the bridge between the release of free energy and the performance of work may be more limited than previously thought. Motions in the myosin strut may be more subtle than was thought.

This result has important implications for the mechanism by which this motor protein transduces chemical energy into mechanical work.

The results from the present research will help in determining 1) a standard protocol to label and isolate skeletal myosin S1 with a single fluorescent probe at any region of the S1 sequence, 2) the route of transmission of force during various conformations of myosin S1, 3) the segmental flexibility in the myosin molecule when constraints are applied, 4) structural and comparative analysis of the flexibility of the strut in various conformations of myosin S1, 5) comparative analysis of the energy required to undergo different conformations, 6) how small modifications to the contractile proteins impact their normal functions and activity, and 6) how dynamic protein structural changes couple with nucleotide states to drive motile processes or lever arm movement.

CHAPTER 2

MATERIALS AND METHODS

Proteins and Reagents

α-Chymotrypsin, trypsin, Staphylococcus aureus V8 protease, ATP, dithiothreiotol (DTT), β-mercaptoethanol, Tris (Sigma, St. Louis, MO), CY3 (Amersham Life Science, Inc., Arlington Heights, IL), phenylmethylsulfonyl fluoride (PMSF), HEPES, DMF (Research Organics, Cleaveland, OH). Potassium iodide (Mallinckrodt, AR) and thallium chloride (Fischer) were procured.

Rabbit skeletal myosin was prepared by the method of Godfrey and Harrington [74]. Rabbit skeletal actin was prepared by the method of Spudich and Watt [75] and stored in G-buffer. Free nucleotides were removed immediately before use by dialysis for 6-8 h in the assay buffer. S1 from rabbit myosin was prepared by digestion of myosin filaments with -chymotrypsin by the method of Weeds and Pope [76]. The protein concentrations were determined using extinction coefficients at 280 nm of 0.55 (mg/ml) 1 cm 1, and 0.75 (mg/ml) 1 cm 1 for myosin, and S1 respectively, and at 290 nm of 0.63 (mg/ml) 1 cm 1 for actin.

Gel Electrophoresis

Laemmli SDS-PAGE(15%) was used to analyze myosin. Fluorescent bands were located in the gels by illumination with a long wave ultraviolet light before staining with Coomassie blue /SYPRO ruby. Densitometric scanning of the gels was performed with a Bio-Rad GS-800 High resolution gel scanner equipped with a computerized integrator. Tricine-SDS-PAGE (16.5%) followed the protocol of Schagger and Jagow as modified by Bio-Rad (Hercules,CA).

CY3–S1 Preparation

CY3 mono-N-hydroxy succinimidyl ester dye (CY3) was dissolved in 50 µl of dimethyl formamide (DMF) and then diluted with 10 mM HEPES buffer (30mM NaCl, pH 8) in a ratio of 1:500.The concentration of CY3 is determined by measuring the absorbance at 550 nm using an extinction coefficient of 150,000 M 1 cm-1. S1 was used for labeling. The volume of CY3 required to be added to S1 is determined for a 1:2, dye:protein, molar ratio and the mixture of CY3-S1 is allowed to react for 10 hours at 4°C. The labeled S1 is dialyzed with 10 mM HEPES buffer (30 mM NaCl, pH 8) for another 10 hours at 4°C, and the labeling ratio was found to have a mean value of 20% for 3 samples. The labeled S1 is dialyzed with 1.5 M ammonium sulphate 50 mM Tris, pH 8, for 10 hours at 4°C. The precipitated protein is removed by centrifuging at 100,000xg (47,000 RPM) in a TLA 100.3 rotor for 1 hour at 4°C. The supernatant containing S1 is purified by hydrophobic interaction FPLC on a phenyl sepharose column (Amersham Biosciences) with 1.5 M ammonium sulphate + 50 mM Tris, pH 8. The major peak containing S1-CY3 eluted with the reverse gradient at 65% buffer.

Limited Tryptic Digestion of S1

Labeled S1 (trypsin/S1 ratio of 1:100 wt/wt for 10 min at 25°C) is digested for 10 minutes and then 0.9 mg/ml of trypsin inhibitor is added to stop the reaction. Tryptic digests were subsequently analyzed by SDS-polyacrylamide gel electrophoresis and visualized for fluorescence on a laser densitometer (Bio-Rad, Hercules,CA). Tryptic digestion of S1 demonstrated that the 50 kDa subdomain was selectively labeled.

Digestion of Myosin with Staphylococcus aureus V8 Protease

Labeled S1 (0.175 mg) was dissolved in 70 µl of 100 mM sodium phosphate (pH

7.8) buffer and denatured with 4 M urea and simultaneously digested with 0.0084 mg of Staphylococcus aureus V8 protease for 40 hrs at 37°C on a shaker. The resulting peptides were analysed on a 16.5% Tris-tricine gels, and the labeled band was cut from the gel and peptide fragments extracted. The peptides were desalted with a G-10 column and lyophilized.

ATPase Assays

ATPase assays were performed by determining the release of inorganic phosphate (Pi) at 25°C by malachite green method. Actin-activated ATPase was assayed in 2 mM MgCl2 , 10 mM imidazole, 1.8 uM CY3-S1, 2 mM ATP with 0-125 μM actin. Rates of ATP hydrolysis were plotted versus actin concentration and fitted.

Mass Spectrometry

Myosin S1-CY3 was concentrated to around 0.25 mg/ml and lyophilized. The lyophilized powder was used for mass spectroscopy analysis. Mass spectra of the fragments were sent for determination (Bio-Synthesis, Inc., Lewisville, TX) with a MALDI-TOF mass spectrometer, positive ion (Applied Biosystems Voyager System 2051) according to the protocols of the manufacturer.

Collisional Quenching

For spectroscopic analysis, 35 μl of 0.1.8 μM CY3-S1 in 0.1 M KCl was diluted 20 fold in 0.1 M KCl, 10 mM imidazole, 2 mM MgCl2, pH 7.0, buffer. Fluorescence quenching was performed with a SLM-Aminco Bowman I I spectrofluorometer (Spectronic Instruments, Rochester,NY). Excitation and emission slits with a bandpass of 8 nm were used for all measurements. Background intensities of samples were negligible in most cases and were subtracted from each sample spectrum to cancel out any

contribution due to the solvent Raman peak and other scattering artifacts. Collisional quenchers of stock solutions 0.3 M KI and 5 mM TlCl were prepared fresh each day in buffer (0.1 M KCl, 10 mM imidazole, 2 mM MgCl2, pH 7). Quenching data was analyzed by fitting to the Stern-Volmer equation [77].

$F_0/F = 1 + K_{SV}[Q] = 1 + k_q\tau_0[Q]$

where Fo and F are the fluorescence intensities in the absence and presence of the quencher, respectively, [Q] is the molar quencher concentration and KSV is the Stern-Volmer quenching constant. The Stern-Volmer quenching constant KSV is equal to *k*q o, where *k*q is the bimolecular quenching constant and o is the lifetime of the fluorophore in the absence of quencher. The Stern-Volmer constants had standard deviations of 10-20% for separate determinations.

Fluorescence Polarization

Samples were prepared as for collisional quenching above. Steady-state polarization measurements were performed with a SLM-Aminco Bowman I I fluorescence spectrometer (Spectronic Instruments, Rochester, NY) equipped with a Glans-Thomson automated polarizers using 1-cm pathlength quartz cuvettes. Excitation and emission slits with a bandpass of 8 nm were used for all measurements. Background intensities of were negligible in most cases and were subtracted from each sample spectrum to cancel out any contribution due to the solvent Raman peak and other scattering artifacts. Polarization values were calculated from the equation by Chen and Bowman 1965 [78].

$$P = \frac{I_{VV} - GI_{VH}}{I_{VV} + GI_{VH}}$$

where I_{VV} and I_{VH} are the measured fluorescence intensities (after appropriate background subtraction) with the excitation polarizer vertically oriented and emission polarizer vertically and horizontally oriented, respectively. *G* is the grating correction factor and is equal to I_{HV}/I_{HH}. All experiments were done with at least 4 sets of samples and averaged values of polarization readings.

Computational Procedures

Conformational Searching, stochastic dynamics simulations, and energy minimizations were performed using Macromodel software (version 8; Schrodinger,Inc). The atomic model of chicken skeletal myosin S1 [20] was reconstructed with the missing surface loops by individual local applications in shells of 1 nm of the Macromodel LOOP algorithm (100 steps) followed by a seeded mixed large scale low mode and Monte Carlo (1000 steps) conformation searches to identify the lowest energy structures. The myosin heavy chain was truncated to residue 792 to leave only the motor domain. Energy minimization and subsequent simulations of the entire modified chicken skeletal myosin S1 was performed using the AMBER* force field and a generalized Born / solvation area model as previously [79]. For both cleft opening and closing simulations, 11 pairs of distance constraints were applied around the 50 kDa cleft. The structure was minimised for a maximum of 5000 iterations with a convergence threshold of 1. The resultant structure is again energetically minimised and is repeated for 10 times and then the resultant structure is dynamically simulated for 100 ps with the maximum number of iterations being 5000 and a convergence threshold of 0.05. Stochastic dynamic simulations allow the atomic model to relax without the constraints of crystal lattice or force restraints.

CHAPTER 3

BIOPHYSICAL AND BIOCHEMICAL TECHNIQUES

Theory of Absorption

In quantum theory, the absorption process can be understood as the transition of a system from one energy state into another by annihilation of a photon [114]. The energy difference ΔE between the two states corresponds then to the photon energy hv, where h is the Planck constant and v is the frequency of the photon. Every molecule has its characteristic absorption spectrum according to its possible energy states and will therefore contribute to absorption within matter. Macroscopically, the decline in radiation intensity by absorption can be described by

$I_T = \hat{I} .e\text{-}p\ \alpha\ (\gamma)$ ----------(1)

where I_T is the transmitted intensity, $\hat{I}$ is the incident, p is the path length through the studied probe and α (γ) is the absorption coefficient. This is Bouguer-Beer's law of adsorption in matter. The absorption, which within solution is also called extinction, is defind as

$A(\gamma)=\log(\hat{I}/I_T)= p.c.\varepsilon(\gamma)$ ----------(2)

where c is the concentration of the absorbing solvent and ε(γ) is the extinction coefficient. This is Beer-Lambert law [114]. Using these equation it is possible to measure concentrations of absorbing substance via light intensities when the path length is known. The above laws neglects scattering processes and thus one should use long wavelengths in experiments (Rayleigh scattering ~1/γ4 to minimize errors).

Theory of Fluorescence

Fluorescence spectroscopy is a type of electromagnetic spectroscopy used for analyzing fluorescent spectra. It involves using a beam of light, usually ultraviolet light, that excites the electrons in molecules of certain compounds and causes them to emit light of a lower energy. Molecules have various energy levels. Fluorescence spectroscopy is primarily concerned with electronic states and vibrational states. Generally, the species being examined will have a ground electronic state of low energy state and an excited electronic state of higher energy. Each of these electronic states has various vibrational states.

Fluorescence spectroscopy has provided a powerful tool for the studies of biological systems. The excitation and emission spectra of intrinsic, extrinsic and prosthetic chromophores can give information on the protein confirmation, ligand binding, and the accessibility of the chromophore to quenchers. After a system in the ground state has absorbed a photon and is now in an electronically excited state, it can relax back to its ground state by either re-emitting a photon or by radiationless decay. The process of re-emission is called luminescence. If the luminescence is due to a quantum theory allowed transition into the ground state this process is referred to as fluorescence, if the transition is not allowed after quantum mechanical rules it is called phosphorescence. The lifetimes of the excited states for phosphorescence is in the order of milliseconds to seconds and can in single cases even reach days. This is in contrast to fluorescence decays which usually occur within 10-7-10-9 s after absorption [80].

The processes and the lifetimes involved in fluorescence are schematically shown

in Fig. 3.1. From the ground state S0, the fluorphore is normally excited into some higher vibrational states of S1, S2 or S3. Immediately thereafter part of the energy is transferred to the surrounding medium as heat and the system thereby relaxes to the lowest vibrational state of S1. This is called internal conversion and it occurs within 10^{-12} s. After solvent relaxation the emission of a photon or radiationless decay from S1 returns the system into one of the vibrational states of S_0 [80]. As can be seen the emitted photons are usually of higher wavelength then the absorbed photons. This effect is termed Stokes shift and with the use of optical interference filters it allows the detection of emitted photons without any contribution of the incident light source, thus leading to a low background and a higher signal to noise ratio in experiments.

The intensity that is absorbed by a fluorescent indicator is just the difference between the incident intensity $\hat{I}$ and the transmitted intensity I_T given by Bouguer-Beer's law. In addition a proportionality factor, termed quantum yield, is introduced which is defined as

$$q=\frac{\text{number of fluorescencephotons}}{\text{number of absorbed photons}}$$

Then the intensity due to fluorescence can be written as

$$I_F=q(\hat{I}-I_T) = q\hat{I}(1-e^{-pc\varepsilon(\gamma)}).\text{----------(3)}$$

In the case of a low indicator concentration (c. $\varepsilon(\gamma)<<1$), the exponential function can be expanded in a Taylor series, where only constant and linear terms contribute significantly to the fluorescence intensity, which can then be approximated as

$$I_F \cong q\hat{I}pc\varepsilon(\gamma).\text{---------- (4)}$$

For low fluorophore concentrations one can also neglect the inner filter effect, where the fluorophore itself reabsorbs part of the emitted fluorescence photons and thus decreases

the fluorescence output. In addition interactions with the solvent or other solutes, i.e, quenching, can attenuate the fluorescence intensity.

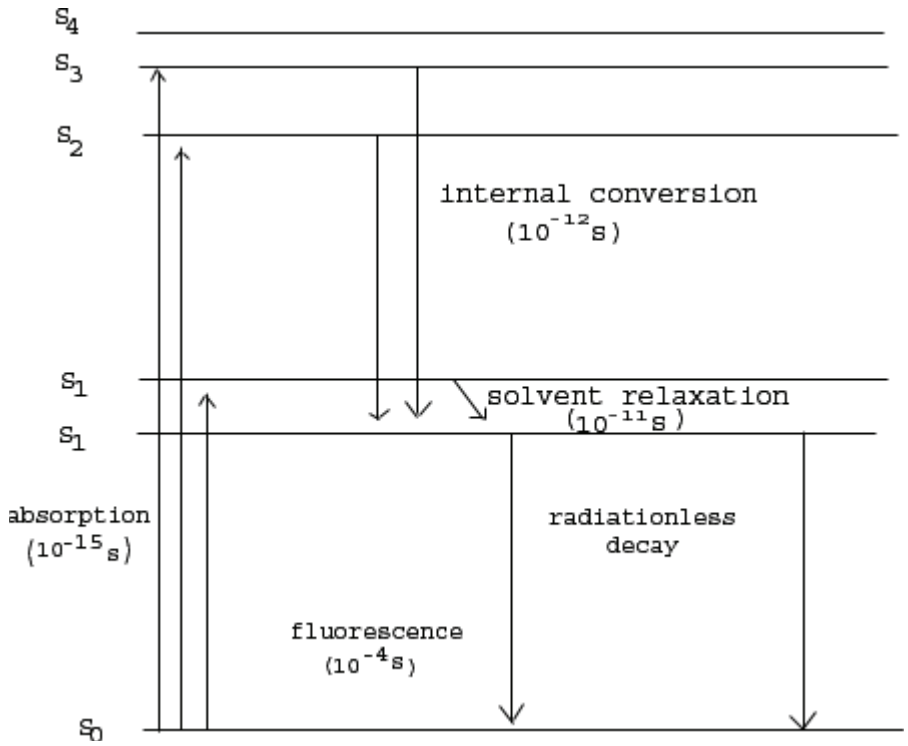

Fig. 3.1: Schematic representation of Fluorescence processes

Analyzing the different frequencies of light emitted in fluorescent spectroscopy, the structure of these different vibrational levels can be determined. The detection of nonfluorescent compounds may be achieved by coupling a fluorescent probe to their primary amino groups. Although fluorescence measurements do not provide detailed structural information, the technique is popular because of its acute sensitivity to changes in the structural and dynamic properties of biomolecules and biomolecular complexes. In fluorescence spectroscopy, the intensity of the emission of the sample is measured. which is represented in Fig. 3.2. The reason for the high sensitivity of fluorescence techniques is that the emission signal is measured above a low background level. Therefore, it is more sensitive than comparing two relatively large signals as in

absorption spectroscopy. The sensitivity of fluorescence techniques is as much as 1000 times more sensitive than absorption spectroscopy.

Fluorescence spectroscopic studies can be carried out at many levels ranging from simple measurement of steady-state emission intensity to quite sophisticated time resolved studies. The information content increases dramatically as various fluorescence observables are time resolved and combined in global analyses of the phenomena of interest. Nonetheless, quite a good deal of information is available from steady-state masurements. Laser-induced fluorescence can be used for a wide array of applications, including the qualitative and quantitative measurements of the concentrations of molecules in a sample. Steady-state fluorometers are used to measure complexation and conformational phenomena of biological molecules. The sensitivity of a spectrofluorometer depends on source intensity, efficiency of the optical system, spectral bandpass of the monochromators, efficiency of the detector. The sensitivity is expressed as the ratio of the signal of a standard sample to the RMS noise level.

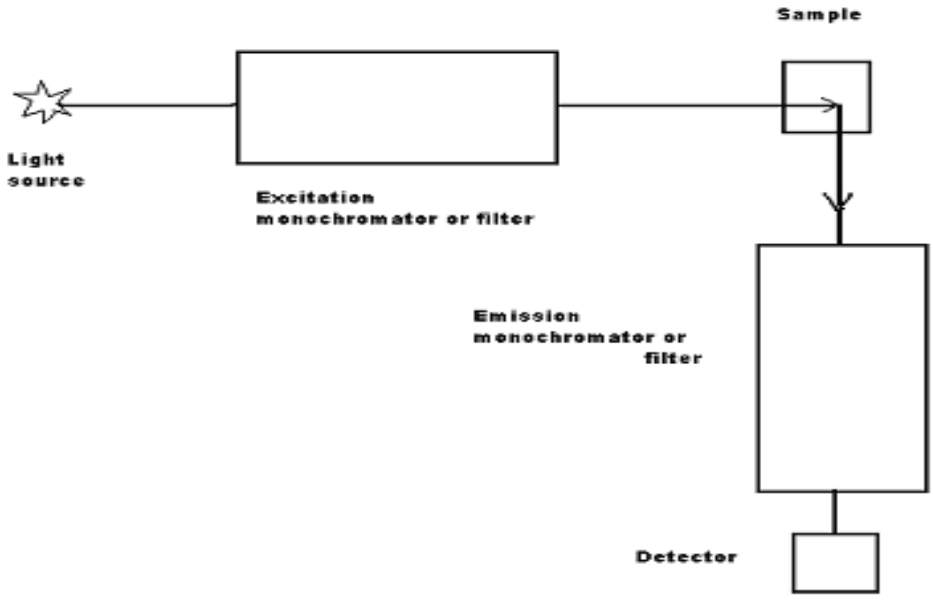

Fig. 3.2: Schematic of a Fluorescence spectrometer

Fluorescence Polarization

This section will discuss the basic theory of fluorescence polarization and how it is used to determine protein dynamics and orientation using extrinsic probes [80]. Fluorescence polarization was first introduced in 1926 by Perrin and later applied by Weber. Fluorescence polarization is a technique to measure molecular orientation and mobility using polarized light and fluorescent label. Light can be described as having characteristics of both particle and wave phenomena. As a wave, light has two components, electric (E) and magnetic (H), which travel in space perpendicular to one another. If polarized light is used to excite a solution of fluorescently labeled molecules, the degree of polarization of emitted fluorescence can be used to measure changes in the environment of the fluor. For example, a small CY3 labeled molecule will tumble rapidly in solution such that emitted fluorescence will be depolarized. However, if this molecule then binds to a larger molecule, it will assume the slower motion of that large molecule and the observed fluorescence will retain a greater degree of polarization. The polarization of the fluorescence can be used to measure the rate of probe rotations and to define the orientations of chromophores. The probability that a molecule absorbs light is maximized for those molecules whose absorption dipoles are aligned along the direction of the polarization of the excited light. The properties of the emitted light are governed by the emission dipole. The electric vector of the emitted light is constrained to lie in a plane defined by the emission dipole and the direction of propagation. These properties allow one to determine the orientation of fluorophores that are rigidly attached to target molecules. When the fluorophores are attached to molecules that are rotating in solution, measurements of fluorescence polarization can be used to determine their rate of rotation.

If linear polarized light is used to excite an ensemble of fluorophores only those fluorophores, aligned with the plane of polarization will be excited. There are 2 possibilities for the emission.

1) If the fluorescence lifetime of the excited fluorescent probe is much longer than the rotational correlation time θ of the molecule it is bound to ($\tau_{fl} >> \theta_{rot}$) (θ is a parameter that describes how fast a molecule tumbles in solution), the molecules will randomize in solution during the process of emission, and, as a result, the emitted light of the fluorescent probe will be depolarized.

2) If the fluorescence lifetime of the fluorophore is much shorter than the rotational correlation time θ ($\tau_{fl} << \theta_{rot}$) the excited molecules will stay aligned during the process of emission and as a result the emission will be polarized.

The fluorescence polarization (P) of a labeled macromolecule depends on the fluorescence lifetime (τ) and the rotational correlation time (Θ)

$(1/P - 1/3) = (1/P_0 - 1/3)(1+\tau/\Theta)$ ----------(5)

where P_0 is the polarization observed in the absence of rotational diffusion. The effect of the molecular weight on the polarization values can be seen from an alternative form of the above equation:

$(1/P - 1/3) = (1/P_0 - 1/3)(1+(kT/\eta V)\tau$ ----------(6)

where k is the Boltzman constant, T is the absolute temperature, η the viscosity and V the molecular volume [80]. The molecular volume of the protein is related to the molecular weight (MW) and the rotational correlation time is given by

$\Theta = \eta V/ kT = \eta\ MW/RT\ (v+h)$----------(7)

where R is the ideal gas constant, v is the specific volume of the protein and h is the

hydration, typically 0.2 g H_2O per gram of protein. Generally, the observed correlation times are about two-fold longer than calculated for an anhydrous sphere due to the effects of hydration and the non-spherical shapes of most proteins.

In the experiment the sample containing the fluorescent probe is excited with linear polarized light and the vertical and horizontal components of the intensity of the emitted light are measured (Fig. 3.3) and the polarization (P) or anisotropy (r) are calculated using the following equations:

Polarization (P) = $(Iv - I_h) / (Iv + I_h)$----------(8)

Anisotropy (r) = $(Iv - I_h) / (Iv + 2I_h)$----------(9)

where Iv is the intensity parallel to the excitation plane and I_h is the emission perpendicular to the excitation plane. They are interchangeable quantities and only differ in their normalization. Polarization P ranges from –0.33 to +0.5 while the range for anisotropy r is –0.25 to +0.4. The maximum values of polarization and anisotropy are possible under any circumstance and indicate that the probe is completely rigid. In contrast, an anisotropy or polarization values of 0 indicates rapid rotation of the probe. Polarization unit is a dimensionless entity and is not dependent on the intensity of the emitted light or on the concentration of the fluorophore [80].

$P = 3 r / 2 + r$ ----------(10)

$r = 2 P / 3 - P$ ----------(11)

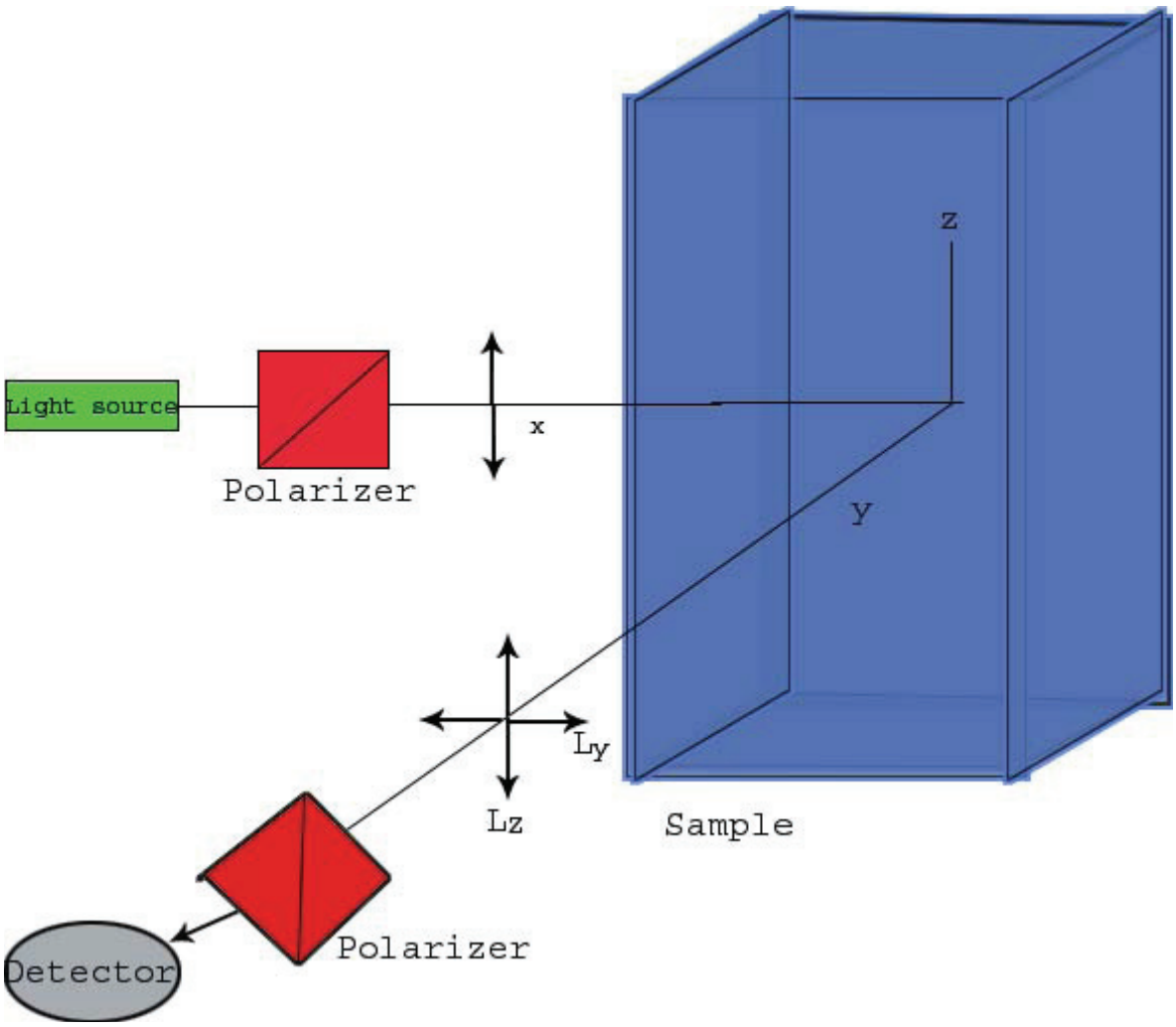

Fig. 3.3: Schematic layout of the general fluorescence polarization experiment

The polarization measurements can give details about any association or binding reactions and any denaturing effects. There are some distinct problems with trying to measure the intrinsic fluorescence of proteins in polarization experiments. Each tryptophan residue has a unique environment with respect to associated amino acids and solvent interactions. Since tryptophan residues are very sensitive, the spectral properties of each residue will be different. Also, tryptophan fluorescence is easily quenched because the indole nucleus tends to donate electrons while it is in the excited state. This makes the interpretation of fluorescent polarization spectra that are based on intrinsic

fluorescence very difficult. The complications resulting from use of the intrinsic fluorescence of proteins in fluorescent polarization spectroscopy has lead to many applications that rely on covalently attached fluorescent labels. If a protein is labeled with a fluorophore, the motions of the fluorophore, are dependent on the nature of the attachment of the fluorophore (e.g rotation around a single teather that links the fluorophore to the protein), the three dimensional shape and motions of the protein (e.g whether the fluorophore is tethered to a sub-unit that can rotate, or if the entire protein tumbles), and on the extent of protein aggregation (e.g tumbling of an aggregation of protein molecules) [81].

Although it is relatively straightforward to directly interpret the fluorescence polarization data in terms of probe angular changes, quantifying the underlying orientational distribution of the labeled proteins themselves requires knowledge of the orientation of the probe relative to the protein. Methods to estimate the local orientations of spectroscopic probes involve assumptions about the hydrodynamic properties of the proteins [82] or prior knowledge of their orientation in a macromolecular system [83].

Quenching

This section discusses the principle of fluorescence quenching. A number of processes can lead to a reduction in fluorescence intensity. Fluorescence quenching is a bimolecular process that reduces the fluorescence quantum yield without changing the fluorescence emission spectrum [80]. It can result from transient excited-state interactions or from formation of non-fluorescent ground-state species.

Quenching may occur by several mechanisms:

1) Collisional or dynamic quenching 2) Static quenching 3) Quenching by energy transfer 4) Charge transfer reaction.

Many other environmental factors influence on fluorescence properties. The three most common are:

1) Solvent polarity which includes interior regions of cells, proteins, membranes and other biomolecular structures.

2) Proximity and concentrations of quenching species.

3) pH of the aqueous medium.

Collisional quenching occurs when the excited fluorophore experiences contact with an atom or molecule that can facilitate non-radiative transitions to the ground state. When quenching occurs by a collisional mechanism, the quenching is an additional process that deactivates the excited state besides radiative emission. Because collisional quenching depopulates the excited state without allowing fluorescence emission, the decrease in fluorescence intensity equates to the decrease in fluorescence lifetime. The dependence of the emission intensity, F on quencher concentration [Q] is given by the Stern-Volmer equation [80].

$Fo/F = To/T = 1 + kqTo[Q]$ ----------(12)

where where T and To is the lifetime in the presence and absence of a quencher, respectively, kq is the bimolecular rate constant for the dynamic reaction of the quencher with the fluorophore. The product of kqTo is referred to as the Stern-Volmer constant or K_{SV}. Collisional quenching is clearly observed when there is a linear decrease in the observed luminescence lifetime with increasing quencher concentration.

Some commonly used quenchers, that are small, aqueous soluble and do not denature proteins by themselves are:

Iodide: Because of its charged nature it efficiently quenches surface residues.

Oxygen: Because of its small size, it can penetrate the protein to some degree.

Acrylamide and nitroxides are neutral.

The accessible fluorophores experience a decrease in fluorescence upon collision with collisional quenchers. Quenching depends strongly on peptide charge. Neighbouring Lys, Arg, His, Asp or Glu residues may greatly affect quenching. If the fluorophores and quenchers are of like charge, then quenching will result in a decrease in kq. Charged quenchers depend strongly on electrostatic effects because quenching depends on ionic strength. In the presence of a quencher, the intensity of the accessible fluorophores decreases according to the Stern-Volmer equation given in equation (13).

$F = Fo,a / (1+K_{SV}[Q]) + Fo,b$ ----------(13)

$dF = Fo - F = Fo,aK_{SV}[Q] / (1+K_{SV}[Q])$ ------(14)

$Fo/dF = 1 / (faK_{SV}) + 1/fa$

where f*a* is the fraction of accessible fluorophores. From this, fa and K_{SV} can be determined from a plot of Fo/dF versus 1/[Q].

Matrix Assisted Laser Desorption and Ionization

This section will discuss the basic theory of matrix assisted laser desorption and ionization (MALDI) and how it is used to determine the molecular mass of peptides [115]. MALDI is a solid-phase ionization technique in which samples are dissolved in a solution containing a chromophoric matrix. The matrix-sample mixture is dried on a target plate and subjected to laser irradiation. Most of the laser energy is absorbed by the matrix, which prevents unwanted fragmentation of the biomolecule. The ionized biomolecules are accelerated in an electric field and enter the flight tube. During the flight in this tube, different molecules are separated according to their mass to charge ratio and reach the detector at different times (Fig. 3.4). In this way each molecule yields a distinct signal. The method is used for detection and characterization of biomolecules, such as proteins, peptides, oligosaccharides and oligonucleotides, with molecular masses between 400 and 350,000 Da. It is a very sensitive method, which allows the detection of low 10^{-15} to 10^{-1} mole quantities of sample with an accuracy of 0.1 - 0.01 %. Protein identification by this technique has the advantage of short measuring time (few minutes) and negligible sample consumption (less than1 pmol) together with additional information on microheterogeneity (e.g. glycosylation) and presence of by-products. The mass accuracy of MALDI-TOF MS will be sufficient to characterize proteins after digesting with enzymes.

The relationship that allows the m/z ratio to be given by

$E = ½ (m/z)v^2$. ----------(14)

In this equation, E is the energy imparted on the charged ions as a result of the voltage that is applied by the instrument and v is the velocity of the ions down the flight path.

Because all of the ions are exposed to the same electric field, all similarly charged ions will have similar energies. Therefore, based on the above equation, ions that have larger mass must have lower velocities and hence will require longer times to reach the detector, thus forming the basis for m/z determination by a mass spectrometer equipped with a time of flight detector [116].

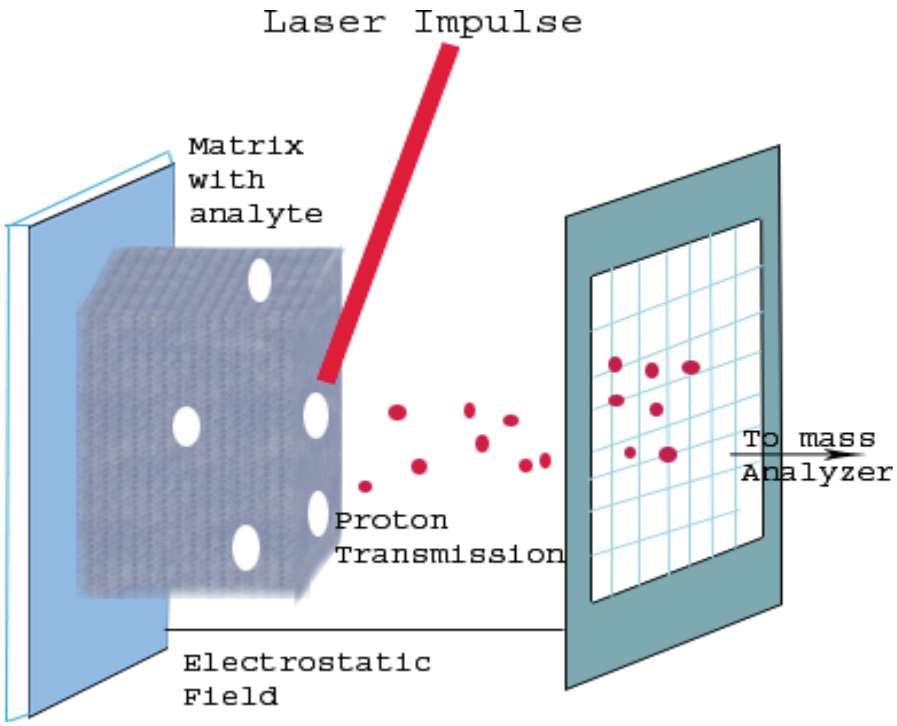

Fig 3.4: Schematic layout of the MALDI-TOF experiment

Hydrophobic Interaction Chromatography

This section will discuss the principle of hydrophobic interaction chromatography (HIC). HIC is a technique for the purification and separation of biomolecules based on differences in their surface hydrophobicity [116]. In HIC the proteins are made to interact with hydrophobic ligands grafted on a hydrophilic matrix. Proteins carry both hydrophilic and hydrophobic areas on their surfaces and at high concentrations of certain salts they precipitate out, allowing interaction with hydrophobic ligands coupled to the chromatographic matrix. This feature allows the use of mild elution conditions to help maintain the biological activity of the sample. HIC media contain ligands that can combine with hydrophobic surfaces of proteins. However, certain salts enhance hydrophobic interactions and adding such salts brings about adsorption to HIC media. To bring about selective desorption, the salt concentration is then lowered gradually and the sample components elute in order of hydrophobicity. Gradients of the salt are formed by mixing two eluents, one containing a high concentration of the salt (buffer A) and one containing a low concentration of this salt (buffer B). But for their salt contents, the two eluents are identical. Chromatography systems usually control the gradient formation by the use of two pumps, one for buffer A and one for buffer B which is depicted in Fig. 3.5.

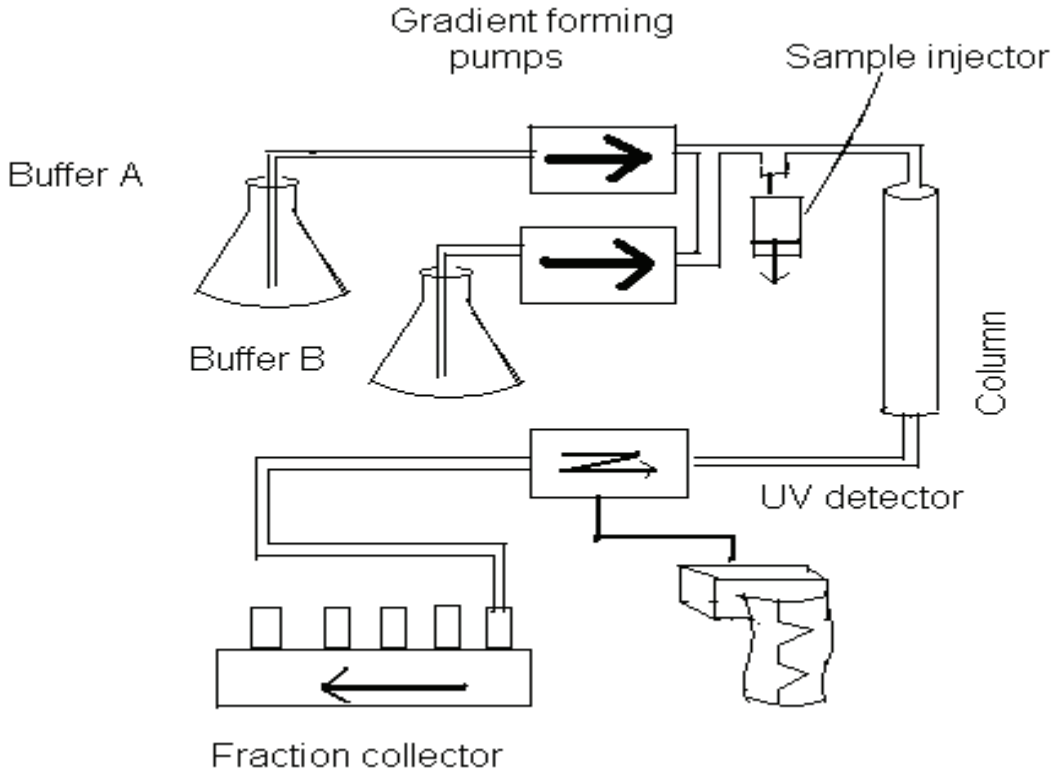

Fig. 3.5: The gradient liquid chromatography (LC) system needs a gradient forming device.

The protein is applied to the column at a salt concentration well above that of the partition zone and is thus adsorbed at the top of the column. When the gradient starts, it will remain fully adsorbed until the salt concentration enters the partition zone. As the salt concentration passes through the partition zone, the protein becomes increasingly desorbed and moves down the column with a velocity reflecting its relative desorption which is depicted in Fig. 3.6.

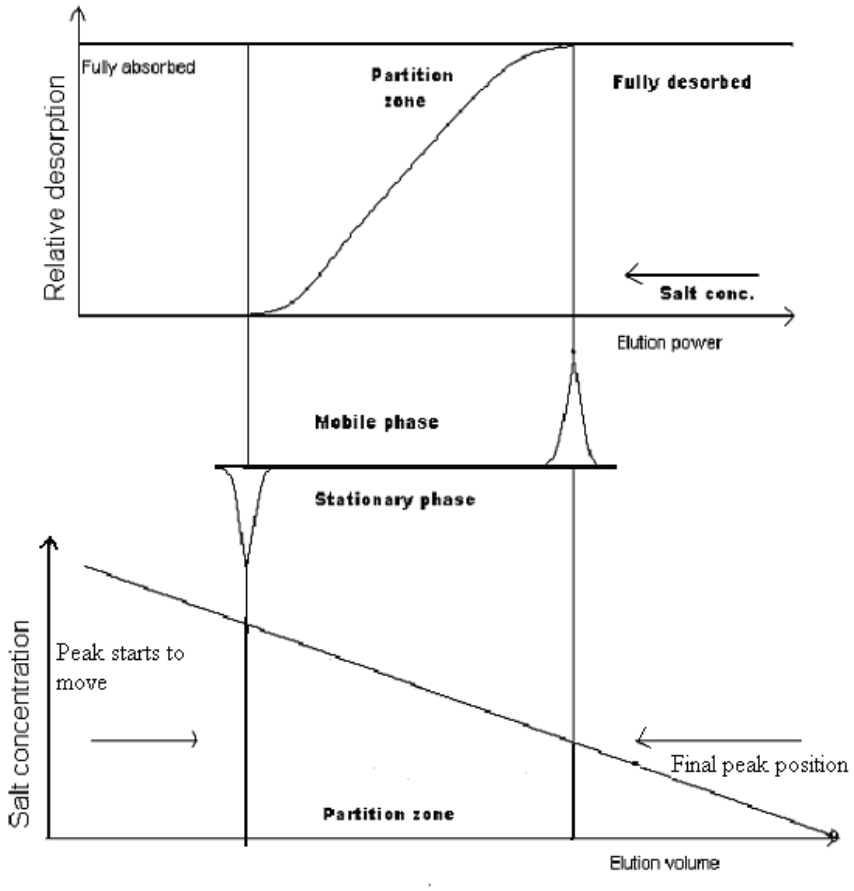

Fig. 3.6: In gradient elution the salt concentration sweeps the partition zone and causes the sample component to accelerate from zero speed to that of the mobile phase. Once at100% desorption the sample component has reached a fixed position in the gradient.

When the salt concentration leaves the partition zone, all protein molecules are fully desorbed. The protein now moves with the same velocity as the eluent and its position within the gradient becomes fixed. The time and thus the travel distance needed to reach this fixed position depends on one hand on the slope of the desorption curve and on the other hand on the slope of the gradient. The possibility to select between ligands of different hydrophobicities is important, since this offers a way to control the salt concentration necessary to adsorb a certain protein. With low-hydrophobicity ligands the difference between adsorption and precipitation may sometimes be so small that certain proteins may partially precipitate under binding conditions. On the other hand, high

hydrophobic proteins may bind sufficiently strongly to high-hydrophobicity ligands and so harsh, potentially denaturing conditions (e.g. organic solvents) are needed to desorb them. Some of the commonly used ligands in order of binding strength are: Ether < Isopropyl < Butyl < Octyl < Phenyl. The difference between the first four ligands is quantitative rather than qualitative. The phenyl ligand, however, offers another type of selectivity, mainly because of its possibility to form additional bonds. By adjusting the salt concentration used to bind the sample molecules, it is possible to control which of the sample proteins that bind and which are just washed through during sample application.

CHAPTER 4

COMPUTATIONAL TECHNIQUES

Energy Minimization

This section discusses the basic theory of energy minimization. Energy minimization is a computational technique which is used to find positions of zero gradient for all atoms, ie, a local energy minimum [119]. In nature, objects tend to change in such a way that their total energy is minimised. The lower the energy, the more stable they are, since they have less capacity to do things. Lower energy states are commonly investigated because of their role in chemical and biological processes. It is generally believed that in many cases a protein's native state possesses the global minimum free energy, or the lowest free energy accessible.

Energy minimization is in many ways the simplest simulation procedure. Energy minimization treats protein molecule as a set of balls with mass connected by rigid rods and springs. Rods and springs have empirically determined force constants so it allows one to treat atomic-scale motions in proteins as classical physics problems. During energy minimization the structure of the protein remains very similar to its crystal structure. The basic idea is that starting from some structure, its potential energy is determined using the potential energy function. Methods which minimize the potential energy are energy minimization techniques (e.g., steepest descent and conjugate gradient).

$$E = E_{bonds} + E_{angle} + E_{dihedral} + E_{non-bonded}$$

$$E_{non-bonded} = E_{electrostatic} + E_{vanderWaals}$$

This function, referred to as a potential function, computes the molecular potential energy as a sum of energy terms that describe the deviation of bond lengths, bond angles and

torsion angles away from equilibrium values, plus terms for non-bonded pairs of atoms describing van der Waals and electrostatic interactions. The potential energy calculated by summing the energies of various interactions is a numerical value for a single conformation [119]. This number can be used to evaluate a particular conformation, but it may not be a useful measure of a conformation because it can be dominated by a few bad interactions. For instance, a large molecule with an excellent conformation for nearly all atoms can have a large overall energy because of a single bad interaction, for instance two atoms too near each other in space and having a huge van der Waals repulsion energy. It is often preferable to carry out energy minimization on a conformation to find the best nearby conformation. Energy minimization is usually performed by gradient optimization where atoms are moved so as to reduce the net forces on them. The minimized structure has small forces on each atom and therefore serves as an excellent starting point for molecular dynamics simulations.

For finding local minima, the conjugate gradient algorithm is used which uses three specific optional parameters: the maximum number of steps, an initial step length used in the search for a minimum and convergence to specify the gradient norm (more precisely the root-mean-square length) at which the minimization should stop. The conjugate gradients method [84] is a more efficient method that combines the information from all previous search directions such that a subsequent search direction is independent to all previous search directions, i.e. the set of search directions forms a linearly independent set. For an n-dimensional vector space, this guarantees convergence in n iterations, as each minimization step reduces the dimensionality of the problem by 1, hence after n iterations, the dimensionality of the problem will be zero, and thus the

minimum has been reached.

Energy minimization methods can precisely locate minimum energy conformations by mathematically homing in on the energy function minima one at a time. The goal of energy minimization is to find a route consisting of variation of the intramolecular degrees of freedom from an initial conformation to the nearest minimum energy conformation using the smallest number of calculations possible. All degrees of freedom in a molecule (3*number of atoms-1) are included in the calculations. Each starting conformation may result in the detection of one minimum energy conformation, although different starting conformations can lead to the same minimum. Energy minimization algorithms measure the energy surface along a series of incremental steps to determine a down-hill direction. The local shape of the energy surface around a given conformation en route to a minimum is often assumed to be quadratic so as to simplify the mathematics. This is illustrated for a one-dimensional slice through an energy surface in Fig. 4.1.

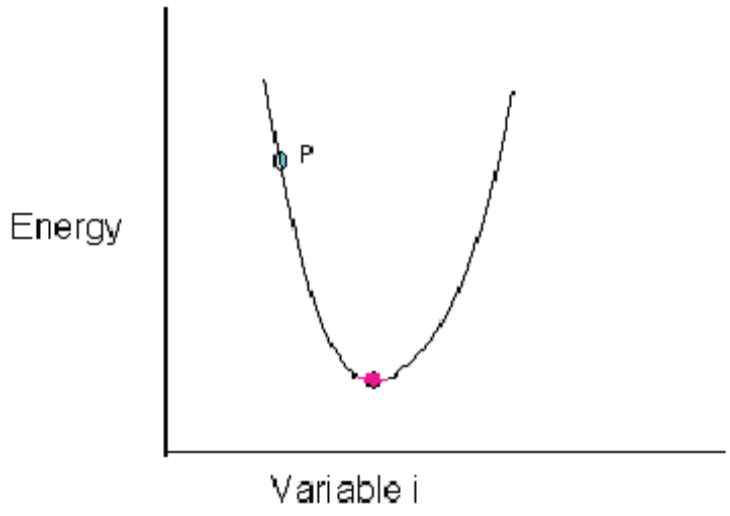

Fig. 4.1: One-dimensional slice through an energy surface

The quadratic energy function approximation is given by a Taylor series [119]:

f(x)=f(P)-b*x+1/2 x*A*x ----------(14)

where P is the current point and x an arbitrary point on the energy surface, b is the gradient at P, and A is the hessian matrix (the second partial derivatives) at P. A and b can be viewed as parameters that fit the idealized quadratic form to the actual energy surface. The A term can improve the fit, allowing larger step sizes toward a minimum, but at the cost of greatly increased computational requirements.

A polyatomic molecule has 3N-6 internal coordinates, where N is the number of atoms. Therefore have to minimize in 3N-6 dimensional space. First derivative of energy with respect to distance, $\delta E/\delta r$, is a force. Instead of one force, we will have a whole array of them: $\delta E/\delta r1$, $\delta E/\delta r2$. This defines a vector which tells us relatively how much and in which sense to alter each dimension in the next step, to go in the direction of steepest descent. The actual size of step taken is reduced as forces get smaller, near the bottom of valley.

The energy landscape of a bio-molecule possesses an enormous number of minima, or conformational sub-states. Nonetheless, the goal of energy minimization is simply to find the local energy minimum, i.e., the bottom of the energy well occupied by the initial conformation ($*$ in Fig. 4.2). The energy at this local minimum may be much higher than the energy of the global minimum.

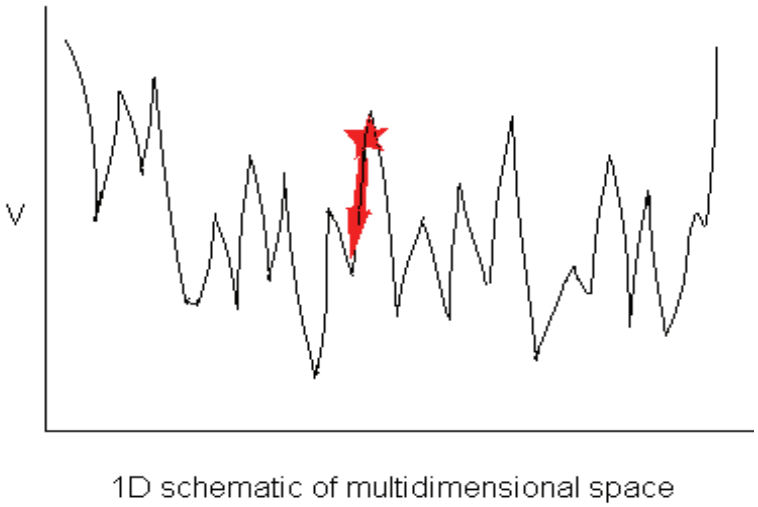

Fig. 4.2: Energy minimization seeks the energy minimum nearest the starting (marked by star) conformation

Uses of Energy Minimization

The energy minimization technique is useful for obtaining a static picture for comparing between states of similar systems.Very often these methods are used if a distorted structure is produced e.g. homology based model. Energy minimization can repair distorted geometries by moving atoms to release internal constraints and make room for a residue, but it will not pass through high energy barriers and stops in a local minima and there is always a local minimum nearby in conformational space. The logical way to overcome local minimum problem is to shake the molecule a bit. Energy minimization can then relieve short inter-atomic distances while maintaining important structural features and removes atomic overlaps in the structure. Stabilizes or reinforces strong hydrogen bonds, breaks weak ones and brings protein to lowest energy in about 1-2 minutes CPU time. Energy minimization can be used to help refine and solve experimental structures obtained from X-ray crystallography and nuclear magnetic resonance.

To perform energy and force calculations of the different conformations, a highly flexible and comprehensive empirical energy function is used which calculates the potential energy for any molecular structure [118]. The potential energy (E) is a summation of several energy terms which can be split up in the following functions:

$$E = E_{bonds} + E_{angle} + E_{dihedral} + E_{non-bonded} \text{ ----------(15)}$$

$$E_{non-bonded} = E_{electrostatic} + E_{vanderWaals} \text{ ----------(16)}$$

The force field description of the interatomic forces is split into two categories: the bonded terms and the non-bonded terms. The bonded terms regroup simple covalent binding as well as the more complex hybridization and π-orbital effects, these are the bonds, angles, dihedrals and improper-dihedrals terms. These terms are schematically drawn in Fig. 4.3. The non-bonded terms describe the Vander Waals forces and the electrostatic interactions between the atoms.

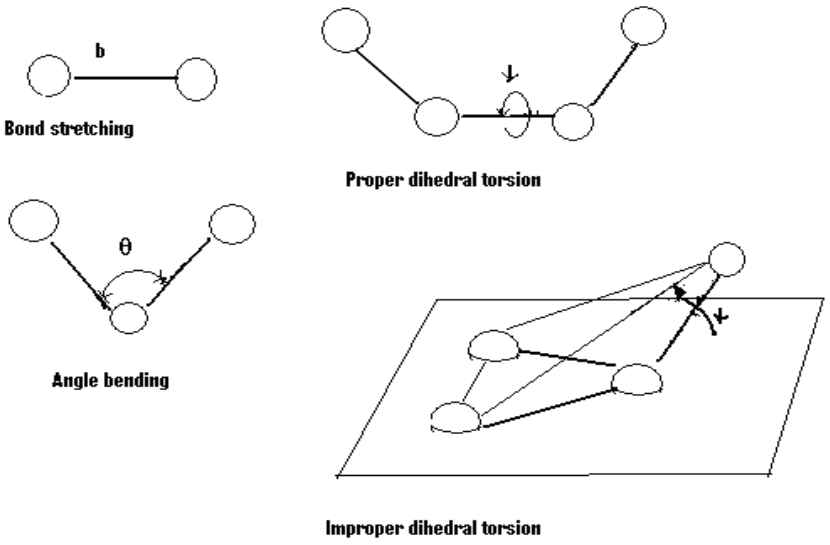

Fig. 4.3: Schematic representation of the bonded interaction terms contributing to the force field: bond stretching, angle bending, proper and improper dihedrals.

Bond stretching

The bond stretching term describes the forces acting between two covalently bonded atoms. The potential is assumed to be harmonic.

$V = k (b - b0)2$ ----------(17)

where b is the distance between the two atoms. Two parameters characterize each bonded interaction: b0 the average distance between them and a force constant k .

Angle bending

The angle bending terms describes the force originating from the deformation of the valence angles between three covalently bonded atoms. The angle bending term is described using a harmonic potential.

$V = k (\theta - \theta 0)2$ ----------(18)

where θ is the angle between the three atoms. Two parameters characterize each angle in the system: the reference angle θ0 and a force constant k.

Torsional terms

The torsional terms are weaker than the bond stretching and angle bending terms. They describe the barriers to rotations existing between four bonded atoms. There are two type of torsional terms: proper and improper dihedrals.

Proper torsional potentials are described by a cosine function.

$V = k [1 + \cos(n\varphi - \delta)]$ ----------(19)

n = 1, 2, 3, 4, 6

where φ is the angle between the planes formed by the first and the last three of the four atoms. Three parameters characterize this interaction. δ sets the minimum energy angle, k is a force constant and n is the periodicity.

The improper dihedral term is designed both to maintain chirality about a tetrahedral heavy atom and to maintain planarity about certain atoms. The potential is described by a harmonic function:

$V = k (\omega - \omega_0)^2$ ----------(20)

where ω is the angle between the plane formed by the central atom and two peripheral atoms and the plane formed by the peripheral atoms.

Van der Waals interactions

Van der Waals interactions and electrostatic interactions are non-bonded interactions, i.e., they act between atoms which are not covalently bonded together. The Van der Waals force acts on atoms in close proximity. It is strongly repulsive at short range and weakly attractive at medium range. The interaction is described by a Lennard-Jones potential [118].

$V = 4\varepsilon [(\sigma/r)^{12} - (\sigma/r)^6$ ----------(21)

where r is the distance between two atoms. It is parameterized by σ: the collision parameter and ε the depth of the potential well. The Lennard-Jones potential is represented in Fig. 4.4.

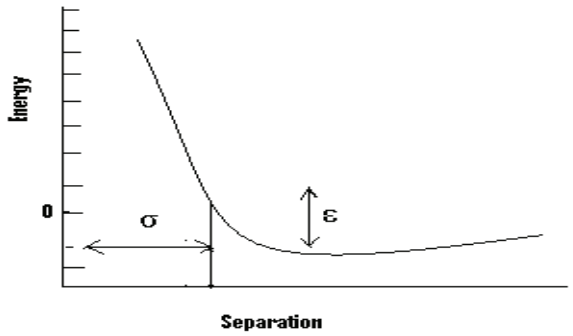

Fig. 4.4: The Lennard-Jones potential. The collision parameter, σ, is shown along with the well depth, ε .

Electrostatic interactions

The long distance electrostatic interaction between two atoms is described by Coulomb's law [115]:

$V = q_1 q_2 / 4\pi\varepsilon_0 r_{12}$ ----------(22)

Where q1 and q2 are the charges of both atoms and r12 is the distance between them. ϵ_0 is the electric susceptibility of vacuum. So finally, the equation for the potential energy describing the force field is:

$$V = \Sigma\, k\,(b-b_0)^2 + \Sigma\, k(\theta-\theta_0)^2 + \underset{\text{dihedrals}}{\Sigma \text{ proper } k\,[1+\cos(n\phi-\delta)]} + \underset{\text{dihedrals}}{\Sigma \text{ improper } k(\omega-\omega_0)^2}$$

$$+\Sigma 4\varepsilon\,[(\sigma/r)^{12} - (\sigma/r)^{6}] + \Sigma q\, q/4\pi\varepsilon_0 r \text{ ----------(23)}$$

Molecular Dynamics Simulations

Molecular dynamics (MD) is a form of computer simulation where atoms and molecules are allowed to interact for a period of time under known laws of physics. Because in general molecular systems consist of a large number of particles, it is impossible to find the properties of such complex systems analytically. MD simulation circumvents this problem by using numerical methods. It represents an interface between laboratory experiments and theory and can be understood as a virtual experiment. MD probes the relationship between molecular structure, movement and function [119]. The main justification of the MD method is that statistical ensemble averages are equal to time averages of the system.

One of the principal tools in the theoretical study of biological molecules is the method of molecular dynamics simulations (MD). This computational method calculates the time dependent behavior of a molecular system. The molecular dynamics method was

first introduced by Alder and Wainwright in the late 1950's. The first protein simulations appeared in 1977 with the simulation of the bovine pancreatic trypsin inhibitor [85]. MD simulations provide detailed information on the fluctuations and conformational changes of proteins and nucleic acids. These methods are used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes.

Biological molecules exhibit a wide range of time scales over which specific processes occur, for example local motions range from 0.01 to 5 Å in a timeframe of 10-^{15} to 10^{-1} s are atomic fluctuations, side chain motions, loop motions. Rigid body motions (1 to 10Å, 10^{-9} to 1s) are helix motions, domain motions (hinge bending) subunit motions. Large-scale motions (> 5Å, 10^{-7} to 10^{4} s) are helix coil transitions, dissociation/association, folding and unfolding**.** Design of a molecular dynamics simulation should account for the available computational power. Simulation size (n=number of particles), timestep and total time duration must be selected so that the calculation can finish within a reasonable time period. However, the simulations should be long enough to be relevant to the time scales of the natural processes being studied. Most scientific publications about the dynamics of proteins use data from simulations spanning nanoseconds to microseconds. To obtain these simulations, several CPU-days to CPU-years are needed.

Molecular dynamics simulations generate information at the microscopic level, including atomic positions and velocities. The conversion of this microscopic information to macroscopic observables such as pressure, energy, heat capacities, etc, requires statistical mechanics. Statistical mechanics is fundamental to the study of biological systems by molecular dynamics simulation. In a molecular dynamics simulation, one

explores the macroscopic properties of a system through microscopic simulations, for example, to calculate changes in the binding free energy of a particular drug candidate, or to examine the energetics and mechanisms of conformational change. The connection between microscopic simulations and macroscopic properties is made via statistical mechanics which provides the rigorous mathematical expressions that relate macroscopic properties to the distribution and motion of the atoms and molecules of the N-body system, molecular dynamics simulations provide the means to solve the equation of motion of the particles and evaluate these mathematical formulas.

The thermodynamic state of a system is usually defined by a small set of parameters, for example, temperature, pressure, and the number of particles. Other thermodynamic properties may be derived from the equations of state and other fundamental thermodynamic equations. The mechanical or microscopic state of a system is defined by the atomic positions, q, and momenta, p, these can also be considered as coordinates in a multidimensional space called phase space. For a system of N particles, this space has 6N dimensions [119]. A molecular dynamics simulations generates a sequence of points in phase space as a function of time, these points belong to the same ensemble, and they correspond to the different conformations of the system and their respective momenta.

Why Apply Molecular Dynamics Simulations to Study Protein Function?

The protein's structure or molecular architecture is sufficient for some of the functions but for most others the function is intimately linked to internal dynamics. In these cases, evolution has optimized and fine-tuned the protein to exhibit exactly that type of dynamics that is essential for its function. Therefore, if we want to understand protein

function, we first need to understand its dynamics. Unfortunately, there are no experimental techniques available to study protein dynamics at the atomic resolution at the physiologically relevant time resolution that can range from seconds or milliseconds down to nanoseconds or even picoseconds. Therefore, computer simulations are employed to numerically simulate protein dynamics. The basic idea behind molecular dynamics is that one determines all forces $F_{i,j}$ between all atom pairs i,j and then moves the atoms in response to them. These forces can arise from covalent bonds, bond angles, torsion angles, Van der Waals interactions, electrostatic interactions, hydrogen bonds, induced dipoles, etc. Using Newton's laws, one can determine the acceleration a from F=ma, and applying this acceleration to the actual speed of the atom at time t will give the speed at time t+dt (in which dt must be very small).

A simple description is that Newton's second law of motion: $F_i=m_i d^2x_i/dt^2$ is solved to find how the position for each atom of the system x_i varies with t. To find the forces on each atom (F_i) the derivative vector (gradient) is calculated. Factors such as the temperature and pressure of the system can be included in the treatment. Molecular dynamics simulation procedures are very popular in the protein field as they have the advantage that they can treat systems where motion is essentially diffusive in character important because of the role of water in protein structure. The procedures can be used to calculate ensemble average properties. It has the ability to calculate free energy differences between slightly different ligands or conformations of a protein.

Advantages of Molecular Dynamics Simulations

Proteins are held together in the native state by hydrophobic interactions, hydrogen bonds and interactions with the surrounding water, whose strength as well as

spatial and temporal distribution affects protein flexibility and hence function. The hydrogen bond framework and the hydrophobic core restrict the flexibility of the peptide chain. Some regions of a protein structure are kept tightly together by a large number of interactions, whereas others, such as loop regions, which often function as binding sites, display enhanced flexibility. Highly ordered regions show little flexibility whereas the large loop appears in different conformations.

Experimental and simulation studies have established that proteins are not rigid molecular objects [86,87] but instead exhibit internal motions that are often essential for their function [88,89]. As a flexible molecule, a protein may populate a large ensemble of different structures. However, modeling protein flexibility is a very difficult task. Experimental laboratory methods such as X-ray crystallography, NMR produce very limited information. A limitation of a crystal structure is that it is a static picture of a single conformation of a protein, obtained in the presence of crystal contact packing forces, and generally under non-physiological conditions which can confound interpretation. What is desired is to augment the insights obtained from x-ray crystallography with non-static analyses. In this regard, classical molecular mechanics simulations offer an extremely powerful tool for analyzing the structural, mechanistic, and energetic properties of biomolecules. Given the atom locations from a crystal structure, one can obtain detailed descriptions of how certain parts of the protein move with respect to time, simulating solvated and physiological conditions. Proteins are dynamics molecules and rarely remain in one confirmation, but must move between different confirmations in order to function. Unfortunately documentation of all the possible conformational structures a protein might exist in is not yet available. The task

of determining the structures of rare protein confirmation is daunting, because it is difficult to isolate proteins in their high energy states for a period long enough to acquire an accurate structure. Since the structures of these high energy confirmations cannot be determined experimentally, a dynamics simulation is next best option for resolving these mysteries.

A protein is a many-body system with 3N degrees of freedom, where N is the number of atoms in the molecule. In most cases the biological interest concerns are only the low frequency motions, which are the essential degrees of freedom [90,91], and typically appear in timescales of nanoseconds to microseconds and play important roles in signaling, activation and enzyme function. Modern computers allow simulations of several nanoseconds, which is long enough to determine the dominant contributions to atomic fluctuations. Since, most of the non-covalent interactions flicker on and off over typically nanoseconds, it is possible to obtain good statistics about structure, molecular interactions from molecular dynamics simulations. Molecular dynamics simulations are often used to gain insights into structure-function relationships in proteins. Characterizing protein flexibility is an important goal to understand the physical chemical principles governing biological function. It is widely accepted that conformational changes of proteins can affect their ability to bind other molecules and that any progress in modeling protein motion and flexibility will contribute to the understanding of key biological functions.

Molecular dynamics methods are useful in refining solution structures of proteins by using constraints derived from nuclear magnetic resonance (NMR) spectroscopy. Therefore, molecular dynamics simulations should provide an attractive means to identify

flexible regions in proteins that can serve as a target for stability enhancement. In particular, loop fragments are often times highly mobile even in generally stable proteins. Such mobile loops are not easily characterized by X-ray crystallography as they may introduce significant disorder in a protein crystal. In fact, partially resolved protein structures are reported in such cases, with the loop fragment missing. However, proposing a single peptide conformation fails to address the mobility of the missing loop. In light of the high variation of loop structures in proteins, one or few conformations may not adequately represent the diversity in the ensemble of conformations assumed by a mobile missing loop. Loops are not the only flexible fragments in a protein. An entire protein can undergo conformational changes that may be essential to its biological function. Compounding evidences from experiment, simulation, and theory indicates that the characterization of protein functions, such as enzymatic reactions, ligand binding, and protein/protein interactions, requires considering a protein native state as a dynamical ensemble of conformations rather than one single structure [92].

While NMR and MD simulations can characterize local backbone fluctuations in the picosecond-nanosecond timescale, slower motions in the millisecond-second range is of crucial interest to many functionally important biological processes and are not well understood. The usefulness of MD simulations of proteins in understanding biological functions is the degree to which the simulations adequately sample the conformational space of the protein. If the timescale is too short to observe many important protein processes, such as slow conformational changes and protein folding/unfolding. It is possible to solve this problem by increasing the simulation time which is computationally demanding and a very slow process.

Molecular dynamics simulations can be extremely useful for studying where biochemical methods cannot be performed. For example, myosin is a large motor protein responsible for muscle contraction. It binds to filamentous actin and uses ATP hydrolysis as an energy source to walk along the filament towards the positive end. Myosin exists in many confirmations, however, only few conformational structures have been determined. However, the beginning and ending confirmations of myosin as it travels along its reaction pathway are well-known. With this knowledge and the proper algorithm, it should be possible to dynamically model myosin as it traverses its reaction pathway. Extensive studies using various techniques indicated that the nucleotide-binding pocket did not experience large conformational changes during the hydrolysis cycle [23,66]. However, the small nucleotide induced conformational changes in the motor domain should be converted into larger movement and these relative changes throughout the structure can be detected using molecular mechanics. The experimental methods can detect few changes at some locations but not minute conformational and structural changes through out the structure. These subtle conformational and structural changes can be detected using molecular mechanics because the nanometer distances separating key functional elements such as the actin binding site, the nucleotide binding site and the lever arm within myosin suggest that the tight packing of the amino acid residues in the motor domain may be influenced to transmit small structural changes from one part of the myosin to another that depends upon structural integrity of many sites along the transmission pathway. Interactions between these sites were examined using molecular mechanics and molecular dynamics simulations in combination with the application of force restraints to determine the overall structural impact of closing and opening the 50

kDa cleft, crosslinking the reactive cysteines in the converter region, and opening and closing the nucleotide cleft.

Disadvantages of Molecular Dynamics Simulations

The obstacles on the route of extending classical molecular dynamics simulations to hundreds of nanoseconds basic time step used in the simulations is limited to resolving the fastest motions in the system. These are often vibrations involving covalent bonds, which occur on the pico-second time scale, and hence require femto-second time steps in the molecular dynamics for sufficient accuracy. Another disadvantage of molecular dynamics simulations is that the collective modes need to be obtained from a predetermined set of protein conformations, usually from a relatively short period of simulations. Such a set usually represents only local fluctuations. A disadvantage of conventional molecular dynamics procedures is that they can only tackle motions with a relatively short time scale, one nanosecond is the approximate upper limit with current computers.

Theory of Molecular Dynamics Simulations

The molecular dynamics simulation method is based on Newton's second law or the equation of motion, F=ma, where F is the force exerted on the particle, m is its mass and a is its acceleration. From a knowledge of the force on each atom, it is possible to determine the acceleration of each atom in the system. Integration of the equations of motion then yields a trajectory that describes the positions, velocities and accelerations of the particles as they vary with time. From this trajectory, the average values of properties can be determined. The method is deterministic, once the positions and velocities of each atom are known, the state of the system can be predicted at any time in the future or the

past. Newton's equation of motion is given by [115]

$F_i = m_i a_i$ ----------(24)

where Fi is the force exerted on particle i, mi is the mass of particle i and a_i is the acceleration of particle i. The force can also be expressed as the gradient of the potential energy,

$F_i = -V_i V$ ----------(25)

Combining equation (24) and (25) yields

$-dv/dri = mi.d^2r_i/dt^2$ ----------(26)

where V is the potential energy of the system. Newton's equation of motion can then relate the derivative of the potential energy to the changes in position as a function of time

$F = m.a = m.dv/dt = m.d^2x/dt^2$----------(27)

Taking the simple case where the acceleration is constant

$a = dv/dt$ ----------(28)

we obtain an expression for the velocity after integration

$v = at + v_0$ ----------(29)

and since

$v = dx/dt$ ----------(30)

we can once again integrate to obtain

$x = v.t + x_0$ ----------(31)

Combining this equation with the expression for the velocity, we obtain the following relation which gives the value of x at time t as a function of the acceleration, a, the initial position, x_0, and the initial velocity, v_0.

$x=a.t^2 + v_0.t + x_0$ ----------(32)

The acceleration is given as the derivative of the potential energy with respect to the position, r,

$a = -1/m \; . \; dE/dr$ ----------(33)

Therefore, to calculate a trajectory, one only needs the initial positions of the atoms, an initial distribution of velocities and the acceleration, which is determined by the gradient of the potential energy function. The equations of motion are deterministic, e.g., the positions and the velocities at time zero determine the positions and velocities at all other times, t. The initial positions can be obtained from experimental structures, such as the x-ray crystal structure of the protein or the solution structure determined by NMR spectroscopy. The initial distribution of velocities are usually determined from a random distribution with the magnitudes conforming to the required temperature and corrected so there is no overall momentum, i.e.,

$P = m_i v_i = 0$ ----------(34)

The velocities, *vi,* are often chosen randomly from a Maxwell-Boltzmann or Gaussian distribution at a given temperature, which gives the probability that an atom i has a velocity v_x in the x direction at a temperature T [115].

$p(v_{ix}) = (mi/2\pi k_B T)^{1/2} \exp[-1/2 \; . \; mivix^2/k_B T]$ ----------(35)

The temperature can be calculated from the velocities using the relation

$T=1/(3N) \; .pi/2mi$ ----------(36)

where N is the number of atoms in the system.

CHAPTER 5

SKELETAL MYOSIN

Importance of Lever-arm Movement

To understand the mechanism for chemo-mechanical transduction, one must decipher the roles of the various domains of the molecule. The crystal structures of myosin and actin, has evolved into the more precise swinging lever-arm hypothesis. This is important to understand how muscle works. Initially, when the cross bridge was discovered it was thought that the whole cross bridge would move in order for the filaments to move. However the x-ray studies could not detect this amount of mass moving. It was not until 1980 that the lever arm hypothesis was formulated, which suggests that instead of the bulky head swing small amount of the tail portion moves relative to the filament axis. The head stays still relative to the actin. The current hypothesis for the mechanism of force generation by actin-myosin-ATP postulates the rotation of the lever arm as the primary mechanical component of the power stroke. The lever arm is an 85 Å long helix, to which the regulatory and essential myosin light chains are bound. During the ATP binding and subsequent ATP hydrolysis, the active site area of S1undergoes subtle conformational changes and as a result the lever arm moves through 11 nm along the actin helix axis (Figs. 5.1 A & B). Now, it seems likely that the myosin-head power stroke works by such a mechanism.

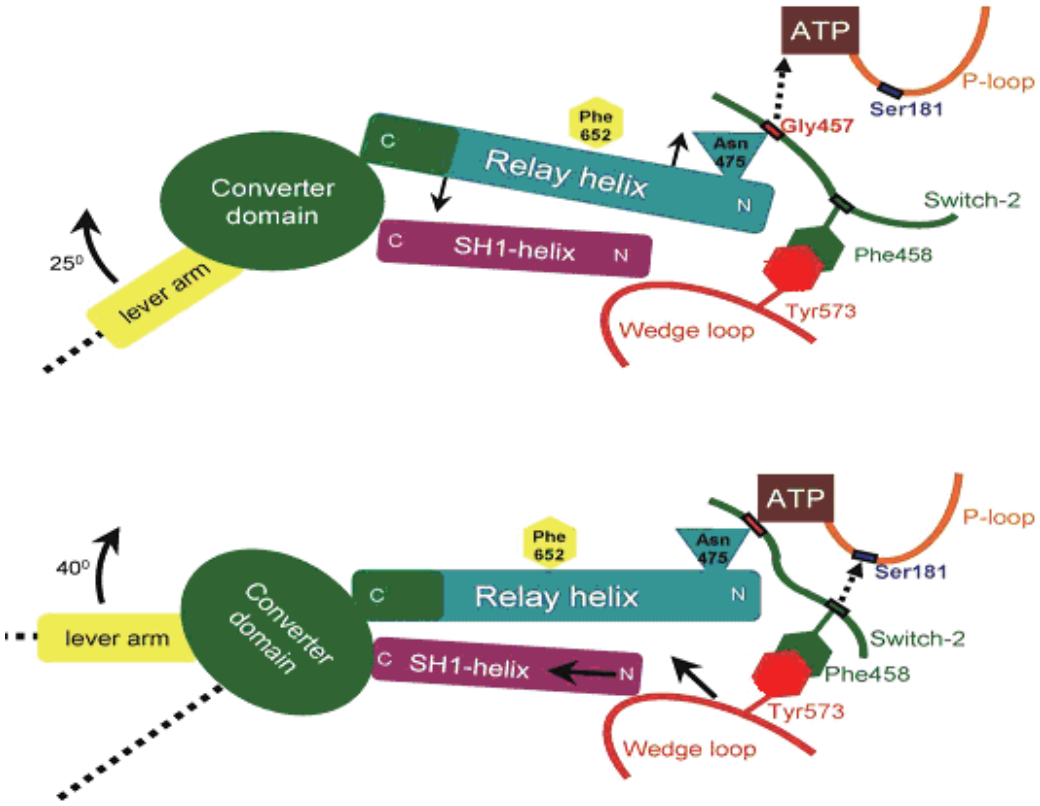

Fig. 5.1A&B: Schematic diagram of the conformational changes of the various subdomains during muscle contraction.

A key feature of myosin motors is the role of the lever arm as an amplifier. The lever arm amplifies small structural changes at the nucleotide binding site to achieve the 110-Å movement along the actin filament that takes place in each ATP hydrolysis cycle. A strong prediction of the mechanism proposed for the movement of myosin along actin is that the length traveled per cycle should depend on the length of this lever arm. Thus, the length of the lever arm should influence the overall rate at which actin moves relative to a collection of myosin heads. This principle has created myosin with extended lever arm which moves actin faster than the native forms. Crystal structures of the myosin II head shows the neck emerging from it, stabilized by two light chains, at an angle that varies by up to 60 degrees depending on whether the head is bound to ATP or to the

products of ATP hydrolysis (adenosine diphosphate and phosphate). So the neck looks like a lever, which led to the idea that it operates as a rigid body to amplify small structural changes in the myosin head, with longer necks leading to a larger displacement. A simple extrapolation to zero speed predicts that the fulcrum point is in the vicinity of the SH1-SH2 region in the catalytic domain.

Recent models of the myosin chemo-mechanical cycle couple movement of this helical neck domain or lever to the opening and closing of the nucleotide-binding cleft and the 50-kDa cleft [35,93]. Removal of one or both light chains from smooth muscle myosin showed a consequent reduction in velocity in vitro of 50–90%. In these studies, the heavy chain ELC and RLC binding domains remained but were devoid of light chains. A striking linear relationship was observed between sliding velocities and the number of light chain binding sites. These data support the model because alteration in the length of the lever arm should alter the size of the mechanical stroke per ATP hydrolyzed and consequently, the velocity of movement. These data are consistent with the movement of a relatively rigid lever arm about a fulcrum point that is very near the reactive thiol region, which is known to undergo structural rearrangement during the ATPase cycle.

In the present dissertation, the movement of lever arm during different conformational states was determined using computational techniques to understand the structural changes in motor domain leading to the movement of lever arm.

Why Study Opening and Closing of 50 kDa Cleft

Structural analysis reveals S1 to be highly asymmetric, consisting of a globular catalytic domain that contains the actin and ATP binding sites and a neck domain made

of the carboxy terminal portion of the heavy chain and two calmodulin like light chains. Within the catalytic domain, the actin and ATP binding sites are on opposite sides of the molecule [20]. These sites had previously been shown to be remote from each other [94,95]. Because the kinetic data indicate that the binding affinity of S1 for actin changes upon binding of ATP and that product release from the active site is activated by binding of actin to S1, a question central to the understanding of chemo-mechanical transduction concerns the communication of these spatially distinct regions. Rayment hypothesized that this communication may be effected through changes in the narrow 50 kDa cleft that subdivides the central 50 kDa domain of the myosin head into upper and lower portions. Potential actin-S1 interface contact sites are also located on both sides of this cleft. In addition, analysis of the active site suggests that the binding site for the γ-phosphate would be located near or at the apex of the 50 kDa cleft [20]. In the absence of nucleotide, the narrow 50 kDa cleft is closed, with actin tightly bound in the rigor complex. Binding of ATP at the active site results in placement of the γ-phosphate at the apex of the 50 kDa cleft, which causes it to open, disrupting the strong binding interaction between myosin and actin. Thus, this cleft provides a mechanism for communication between the spatially distinct actin and ATP-binding sites. Hence, it is possible that changes at one site may be transmitted through this 50 kDa cleft to the other site and vice versa, providing a pathway for the transduction of this information. Hydrolysis of the bound ATP to ADP and Pi results in changes in the 50 kDa cleft that are transmitted via the reactive thiol region of the 20 kDa domain to the light chain binding domain, which acts like a lever arm and swings 5 nm or more relative to the actin binding site into a cocked configuration. These proposed changes are thought to occur

during or just after the hydrolysis of ATP. Following hydrolysis and priming of the lever arm, rebinding of actin causes further movement of the lower domain of the 50 kDa segment relative to the upper domain, which results in release of Pi from the base of the nucleotide-binding cleft and subsequently the release of ADP. The release of the nucleotide products results in complete closure of the 50 kDa cleft [93] which is represented in Fig. 5.2. The closure of the cleft leads to conformational changes in the 20 kDa reactive thiol domain that are postulated to cause the light chain binding domain to rotate, resulting in a power stroke of 5 nm or more. Thus the rigor actin-S1 complex is reformed.

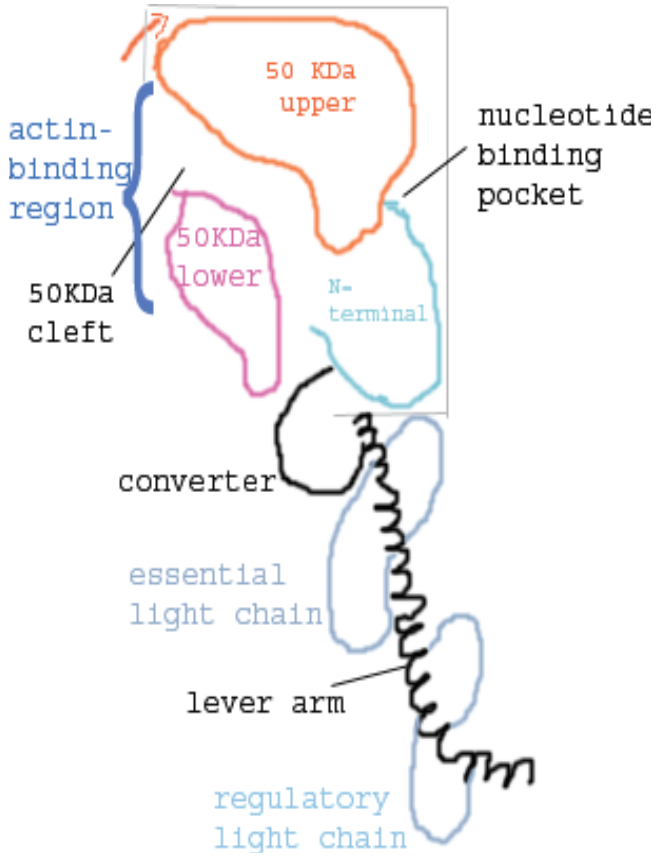

Fig. 5.2: Schematic representation of the different subdomains in the Scallop S1–ADP structure. Arrow indicates rotation of the 50-kDa upper subdomain, which closes the 50-kDa cleft when myosin binds strongly to actin

More support for the central role of this cleft comes from the crystal structures of the Dictyostelium motor domain in the presence of either beryllium or aluminium metallofluoride magnesium-ADP complexes [93]. The beryllium fluoride complex mimics the ATP-bound/prehydrolysis state, whereas the aluminum fluoride complex seems to be closer to the ADP-Pi or post-hydrolysis state. These two crystal structures indicated that hydrolysis of ATP induces a change in the cleft that may be transmitted to the actin-binding site and possibly the neck region of the molecule as well. Structural studies have implicated the 50 kDa cleft both in communication between the spatially distinct ATP and actin-binding sites and in transmitting changes to the lever arm [35]. Mutagenic analysis of the residues that comprise this cleft confirm its importance in the actin-myosin chemomechanical cycle. Analysis of this cleft reveals that there is a high proportion of evolutionarily conserved residues making up its upper and lower borders, again supporting the idea that this region is pivotal to myosin function. Random mutagenesis of the region of Dictyostelium myosin that forms the apex and the lower border of the cleft resulted in distinct phenotypes in vivo and in vitro [20]. The defect in ATP hydrolysis exhibited by several of the more severe mutations suggests that this cleft must be able to alter its conformations for hydrolysis to take place. Further analysis of these mutated myosins as well as mutagenesis of other residues that line this cleft will lead to greater insights into the roles of this region in chemo-mechanical transduction. In the present dissertation, residues around the 50 kDa cleft were selected to open/close the 50 kDa cleft and determine changes in myosin S1 and communication between spatially distinct regions.

Nucleotide Binding Site

The nucleotide binding site is located at the interface between the 50 kDa and the N-terminal subdomains. It is composed of three loops that are conserved not only among motor proteins but also among the G-proteins. These are the P-loop that is a common feature of a large number of enzymes that bind nucleotide, the switch-1 loop and the switch -2 loop [96,97], together they form the so-called phosphate tube. The switch-1 loop and the switch-2 loops are located in the upper and lower 50 kDa domains, whereas the P-loop belongs to the N-terminal 25 kDa domain. The two switch loops adopt different confirmation, and serve as a switch for information transduction by changing their confirmation. The position of the switch-1 and switch-2 loops can be used to classify different, conformational states of myosin. Although crystal structures have been obtained for a large variety of G-proteins in all stages of the hydrolysis cycle, in myosin the structural picture is incomplete, as only the weakly actin-bound states have been captured. This may be a consequence of the complexity of myosin active sites relative to those of G-proteins. Conformations corresponding to switch-1 closed/switch-2 closed (C/C), switch-1 closed /switch-2 open (C/O), and switch-1 open/switch-2 open (O/O) have been identified by crystallography. A key structural part of the nucleotide induced conformational changes in the core of the motor domain is the switch-2 element, which consists of the switch-2 helix and the switch-2 loop.

The ATP binding pocket of myosin not only functions to bind and hydrolyze ATP but, importantly, it governs the release of actin from the rigor complex upon nucleotide binding and the reciprocal release of nucleotide products in response to rebinding of actin. Furthermore, the characteristics of the active site must account for the relative

stability of the myosin-ADP-Pi complex in the absence of actin. Structural rearrangements of the myosin upper-50 kDa subdomain are thought to play a key role in coordinating actin binding with nucleotide hydrolysis during the myosin ATPase cycle. Such rearrangements could open and close the active site in opposition to the actinbinding cleft, helping explain the opposing affinities of myosin for actin and nucleotide.Changes in the conformation of the active site of myosin subfragment-1 (S1) may be linked to the production of force during the power stroke.

Although domain movements of the lever arm and actin-binding cleft are now well established in myosin, conformational changes at the active site have proven difficult to observe by crystallography. The active site would be expected to open in the strongly actin-bound states to allow for the exchange of ADP and ATP, and close upon ATP binding to provide a catalytically competent conformation for hydrolysis. Rayment predicted from their original myosin crystal structure that opening of the active site could be triggered by actin-binding cleft closure [35]. Further, pivoting the upper-50-kDa subdomain around a hinge located near the active site allows for greatly improved fitting of the skeletal myosin crystal structure to cryo-EM reconstructions of the acto-myosin complex [98]. Such a pivot would close the actin-binding cleft while opening the active site, as predicted by Rayment et al [35].

To understand the structure-function relationship of this ATPase-dependent actinmyosin interaction, Dictyostelium myosin II has been extensively used for site-directed mutagenesis. By generating a large number of mutant myosins, two hydrophobic actinbinding sites have been revealed, located at the tip of the upper and lower 50 kDa subdomains of Dictyostelium myosin, one of which is the cardiomyopathy loop.

Furthermore, the slight change in relative orientation of these two hydrophobic sites around the strut loop has been shown to work as a switch to turn on and off the strong binding to actin. Once the switch is turned off, myosin enters in the weak-binding state, where ionic interactions between actin and the loop 2 of myosin become the dominant force to maintain the actin–myosin association. The details of actin-myosin interactions revealed by the Dictyostelium system can serve as a framework for further examinations of the myosin superfamily proteins. In the present dissertation, the nucleotide binding site is examined in various conformational states of myosin to pin-point the structural rearrangements of myosin S1.

Cross linking of SH1-SH2

The myosin head is represented by four subdomains, which are connected together by three joints serving as key structural elements [99]. These three joints coordinate the conformational changes that occur in the myosin head as it hydrolyzes ATP. One of these joints, located near the converter region of myosin, is a bent -helix, which contains the reactive sulfhydryl groups, SH1 (Cys-707) and SH2 (Cys-697). These sulfhydryls Cys-707 and Cys-697, are found on consecutive regions of α helix joined by a short turn, and spatially separated in the chicken myosin structure by ~20 Å [20]. SH1 and SH2 groups can be crosslinked by reagents with widely varying crosslinking spans (5 Å to 12-14 Å), and even by disulfide bond formation, and that binding of nucleotides to S1 promotes such crosslinking [100]. Crosslinking of SH1 and SH2 results in the trapping of nucleotide, in ADP form, at the active site of the enzyme [101]. This helix appeared also to be functionally important as the ATPase activity of S1 was inactivated by its crosslinking.

Clearly, some structural alteration must occur along with hydrolysis that shortens the distance from SH-1 to SH-2 which has not yet been reported for the crystal structures. Previous studies, using a method that isolated and measured the rate of the SH1-SH2 crosslinking step, showed that this helix undergoes ligand-induced conformational changes. Early crosslinking studies demonstrated that when nucleotides (MgADP or MgATPγS) are bound to subfragment-1 (S1), the rate of SH1-SH2 crosslinking is increased [102]. Crystallographic evidence for the nucleotide induced destabilization of the SH1-SH2 helix has been obtained so far only from the scallop S1.ADP structure [41]. In this structure, the electron density of the SH1-SH2 region could not be defined, suggesting that the helix is disordered and that its order-disorder transitions may be functionally important. In fact, the SH1-SH2 helix has been shown in several previous studies to impact the function of myosin. Myosins with glycine residues in the helix mutated to alanines have altered ATPase activities and show a complete loss of the motor function [103,104]. Similarly, the modification of either SH1 or SH2 has resulted in the loss of myosin function [105]. These results suggest that the SH1-SH2 helix undergoes conformational changes and that the flexibility of this helix may be important for the proper function of myosin.

If we assume that conformational changes in the SH1-SH2 helix are involved in lever arm movement, then the crosslinking of SH1 and SH2 groups or modification of SH1/SH2 can disrupt the mechanical function of S1 by altering the flexibility of this helix or its coupling to the lever arm. The SH1-SH2 helix is transformed from a highly flexible structure in the ATP state (S1.ADP) into a rigid one in the prepower stroke state (S1.ADP.Vi) [99]. Therefore, In principle, crosslinking experiments should provide

insight into the conformational dynamics of this helix in the transition states of S1 and ATP in solution. The molecular dynamics simulations should provide important information about the conformation and crosslinking of the SH1-SH2 helix in the transition states of S1. The flexibility of the SH1-SH2 helix in the various transition states could lead to multiple orientations of the lever arm, which was previously observed in cryo-electron microscopy and electron paramagnetic resonance studies [106,107].

Other indications of functional and structural importance of this region came from studies on the selective modification of the SH1 group on S1. It was shown that this modification affects strongly the Mg-ATP hydrolysis cycle of S1, mostly by accelerating the release steps of ATP hydrolysis products. The activation of Mg-ATPase activity of S1 by actin was also altered greatly by SH1 modification [108]. The motor function of myosin heads in the in vitro motility assays was blocked by SH1 modification [105,109]. SH1 modification enhanced the ability of S1 to activate regulated actin. The conformation of S1 alone and in complexes with nucleotides was also altered by SH1 modification [110]. These data show that the SH1-SH2 helix is a functionally important site on S1, but the mechanism by which this helix is involved in the force generation cycle is still unknown. Atomic resolution structures of chicken skeletal S1 and of Dictyostelium myosin S1 complexes with nucleotides, and nucleotide and phosphate analogs confirmed the indications of many solution studies on the key role of the SH1-SH2 helix in myosin function [71,20]. In particular, the transition between the Mg-ADP·BeFx and Mg-ADP·Vi structures of Dictyostelium myosin S1, which models the transition between the Mg-ATP and Mg-ADP·Pi complexes of S1, involved changes in the switch II loop, the lower 50 kDa region, and in the orientation of the SH2 group and,

consequently, the pivoting of the SH1-SH2 helix [71,73]. The location of the SH2 in the atomic structure of S1 and the above observed changes in SH2 environment in the Mg-ADP·BeFx and Mg-ADP·Vi complexes of Dictyostelium myosin S1 suggest its involvement in signal transduction on S1 from the ATP and actin sites to the mechanically important lever-arm region. In the present dissertation crosslinking of SH1-SH2 in various conformational states were analyzed to determine the structural changes in myosin S1.

Why Analyze the Changes at the Strut?

The actin binding site is split by a deep cleft that closes on strong binding to actin. The cleft is straddled by a short polypeptide known as the strut. Strut forms one of the two elements of protein chain joining the upper and lower 50 kDa domain and may control cleft closure. The upper and lower 50 kDa subdomains, separated by the 50 kDa cleft are connected to each other by three loops, the switch II loop, loop 2, and a short loop of four residues (Asp_{601}, Pro_{602}, Leu_{603}, and Glu_{604}) which is known as the strut loop. The switch II loop is at the bottom of the 50 kDa cleft, whereas the other two loops are at the distal end of the cleft. The strut loop has a stretched conformation, encompassing two α-helices in the upper and lower 50 kDa subdomains. Although the peptide backbone of the loop is exposed to the solvent, the side chains of Leu603 and Gln604 are in contact with other hydrophobic side chains in the upper 50 kDa subdomain, whereas those of Asp601 and Pro602 are somehow exposed. Therefore, it is expected to play a role as a strut and keep the relative disposition of the two subdomains. The upper 50 kDa domain moves essentially as a solid body on binding to actin. Furthermore, this movement opens

the nucleotide binding pocket by moving switch 1. The upper 50 kDa domains is tethered by loop 1 and loop 2 and supported by the strut. The sequence and length of the strut loop are strongly conserved among all myosin family members, therefore the strut loop has functional importance.

Previous studies on *Dictyostelium* myosin II reported about the functional roles of the strut loop, especially in a transition between the strong and weak binding states during the ATPase cycle. When a single residue was inserted into or deleted from the strut loop of Dictyostelium myosin II. These mutant myosins and their fragments showed that 95 especially the insertion mutations, selectively abolished the strong-binding state. Insertion or deletion of a single residue in the strut sequence abolished in vivo motor functions of myosin. Unlike these insertion or deletion mutants, the Ala mutant completely complemented the phenotypic defects of myosin-null cells, showing that although the residues in the strut sequence are conserved among various types of myosins, their side chains are not essential for maintaining in vivo functions of myosin. These results imply that the particular length, and not the sequence, of the strut loop is critical for the motor functions of myosin.

In vitro studies of purified full-length myosins and their subfragment-1s revealed that the insertion mutants virtually lost the strong binding to actin although their motor functions in the absence of actin remained almost normal, showing that only the hydrophobic actin-myosin association was selectively affected by the insertion mutations. Unlike the insertion mutants, the deletion mutant showed defects both in the strong-binding state and the rate-limiting step of ATPase cycle. These results indicate the functional importance of the strut loop in establishing the strong-binding state of myosin

and thereby achieving successful power strokes. In the present dissertation a residue in the strut is labeled with CY3 and the structural changes in the strut will be analyzed in various conformations of myosin.

CHAPTER 6

EXPERIMENTAL RESULTS

Fluorescence Labeling

The labeling methodology developed herein has potentially general applicability. By using various combinations of molar ratios of protein and different fluorescent labels, it is possible to isolate singly labeled protein of any region of interest and study the conformational and structural changes of the labeled region of a protein and hence gain more understanding about the mechanism of mutated proteins. In the present research, CY3 N-hydroxysuccinimidyl monoester (CY3) is used as a hydrophobic tag and myosin subfragment-1 (S1) as the target protein, S1 is reacted with a half molar ratio of CY3 to minimize the possibility of dual labeling of the S1.The labeled S1 is then separated from unlabeled S1 by hydrophobic interaction chromatography. Then the labeling site is mapped and the fluorescence is analyzed (Fig. 6.1). The attachment of CY3 to a surface residue perturbs the hydrophobicity of S1 sufficiently that it can be separated by hydrophobic interaction chromatography from unlabeled S1.This hydrophobic tag can, in principle, convey different retention times to alternative labeling sites such as heavy chain versus light chain (Fig. 6.2). Multiple site specific labeled products are detected in a single hydrophobic interaction chromatography run, although the additional peaks were not analyzed in the present research.

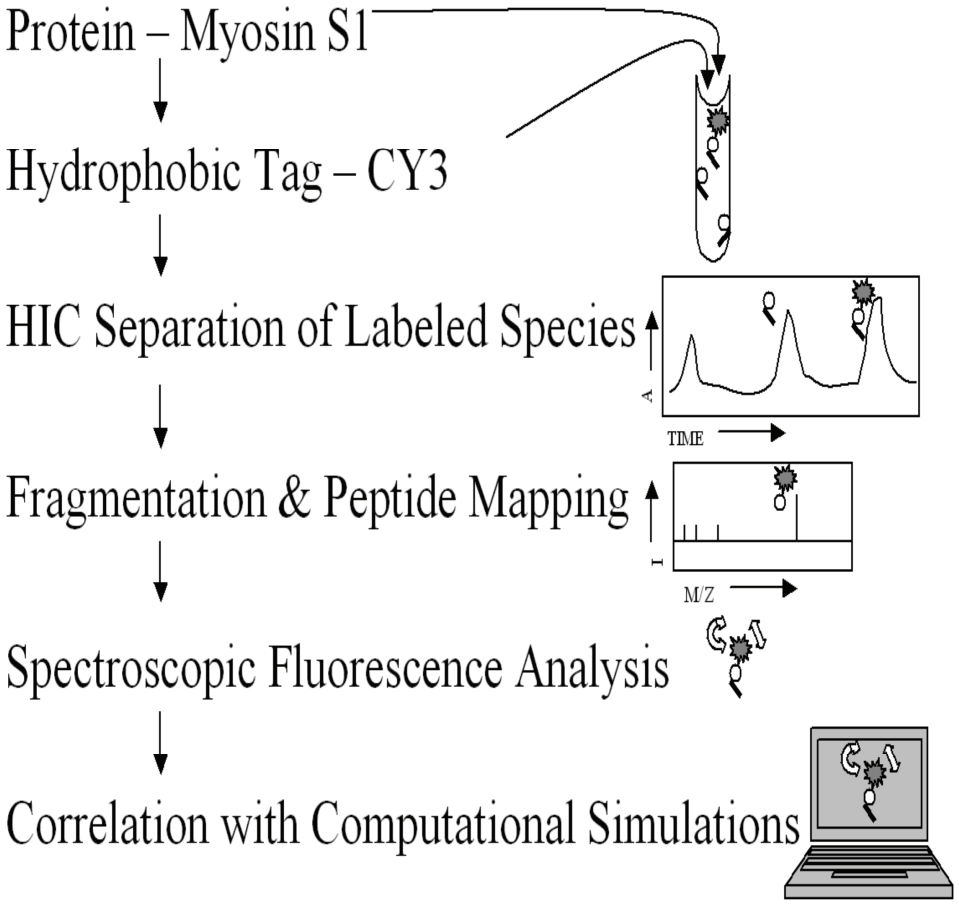

Fig. 6.1: Schematic diagram of the research : Using CY3 N-hydroxysuccinimidyl monoester (CY3) as a hydrophobic tag and myosin subfragment-1 (S1) as the target protein, S1 is reacted with a half molar ratio of CY3. The labeled S1 is then separated from unlabeled S1 by hydrophobic interaction chromatography. Then the labeling site is mapped and the fluorescence is analyzed. These results will be correlated with computer simulations

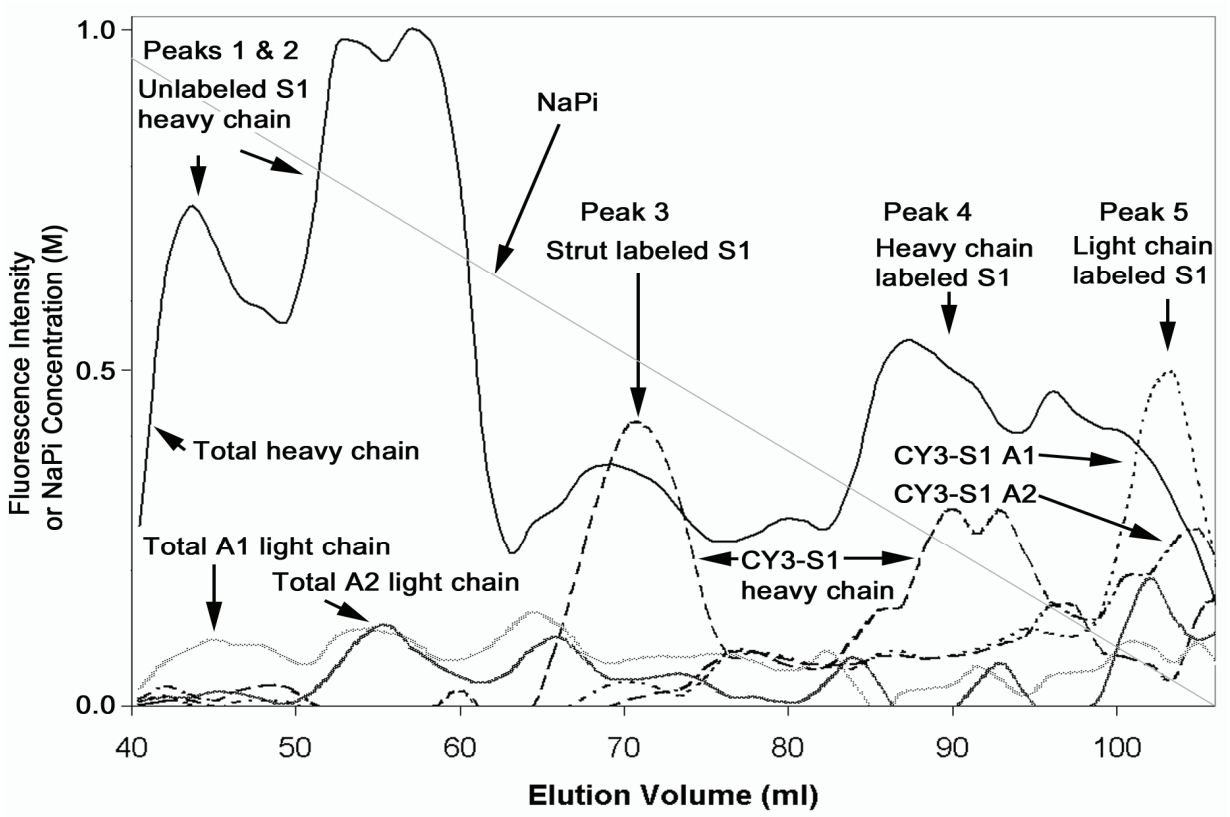

Fig.6.2 : Chromatograph of hydrophobically tagged CY3-S1 purification. The labeled S1 is separated from unlabeled S1 by hydrophobic interaction chromatography. The retrieved fractions were analyzed by SDS-PAGE and fluorescence densitometry first for CY3-labeled protein and again after SYPRO-ruby staining for total protein. Unlabeled S1 was less hydrophobic and eluted earliest from the column in two peaks. The A1 isoform was partially separated from the A2 isoform, The CY3-labeled S1 was more hydrophobic and eluted later in at least 3 peaks of which two were labeled on the heavy chain and one on the light chains. The first labeled peak (designated peak 3) was best resolved and was chosen for further peptide mapping. The integrated intensities of total protein from the labeled peaks were 46% of the overall recovered protein.

The labeling site in the major peak was mapped by tryptic digestion to the 50 kDa subdomain of S1 and the specificity of heavy chain labeling was greater than 95% (Fig. 6.3A). A total proteolytic digest of denatured CY3-S1 with V8 protease was separated by tricine gel electrophoresis (Fig. 6.3B). The approximately 2 kDa labeled band was isolated and further analyzed by MALDI-TOF to yield a precise mass of 2.083 kDa (Fig. 6.3C) that corresponded exactly with the predicted mass of a single proteolytic fragment

from the 50 kDa subdomain of S1 and no other labeled peptide mass was detected (Table 1). This peptide contains two nearby lysine residues, K598 and K600, that are localized to the strut sequence of myosin. The high specificity of the N-hydroxysuccinimidyl ester reaction for primary amines dictates that one of these two lysines must be labeled (Fig 6.3D).

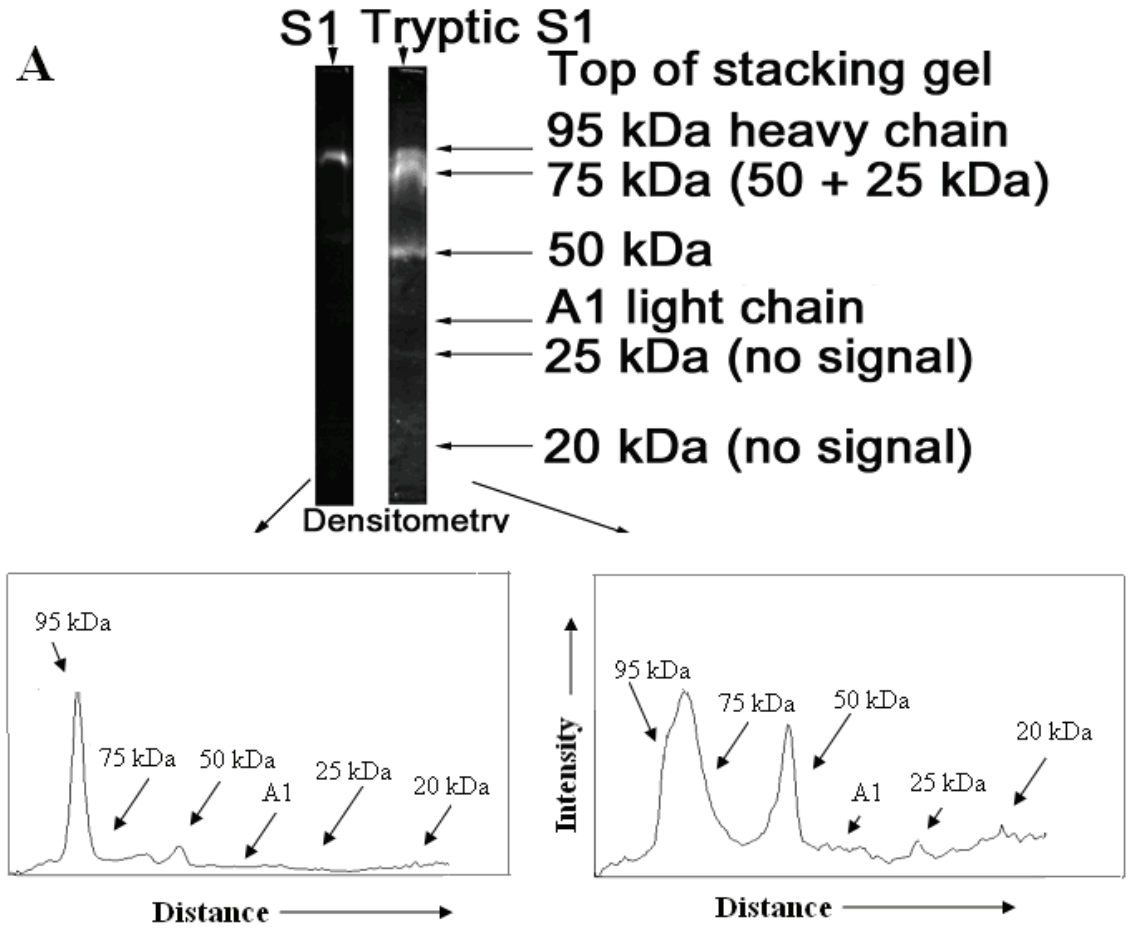

Fig. 6.3: Peptide mapping of CY3-S1. (**A**) The CY3-S1 from peak 3 was analyzed by partial tryptic digestion. Fluorescence densitometry of the SDS-PAGE from CY3-S1 before and after tryptic digestion indicated that at least 95% of the CY3 fluorescence was localized to the 50 kDa subfragment of the heavy chain. The labeled tryptic subfragments of the heavy chain were the 75 kDa (which contains the 50 kDa and 25 kDa subfragments) and the 50 kDa subfragments but not the 20 kDa and 25 kDa subfragments.

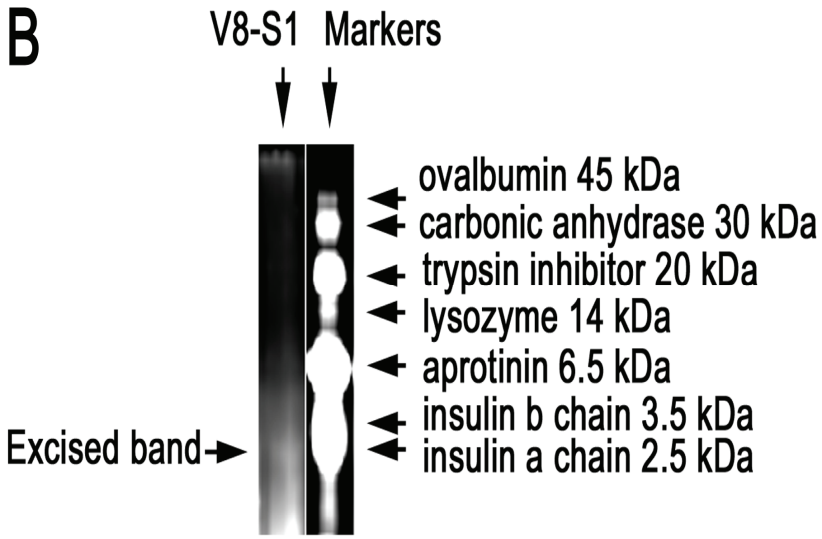

Fig 6.3B: CY3-S1 was digested under denaturing conditions with V8 protease to cleave specifically after aspartate and glutamate residues. This partial digest yielded a single 2 kDa fluorescent fragment on a tricine peptide gel. CY3-labeled fragments must be greater than 1.1 kDa according to sequence analysis. (See Table 1.) Some autofluorescence is detected near the buffer front, so for precise analysis MALD-TOF was used to identify small mass peptides.

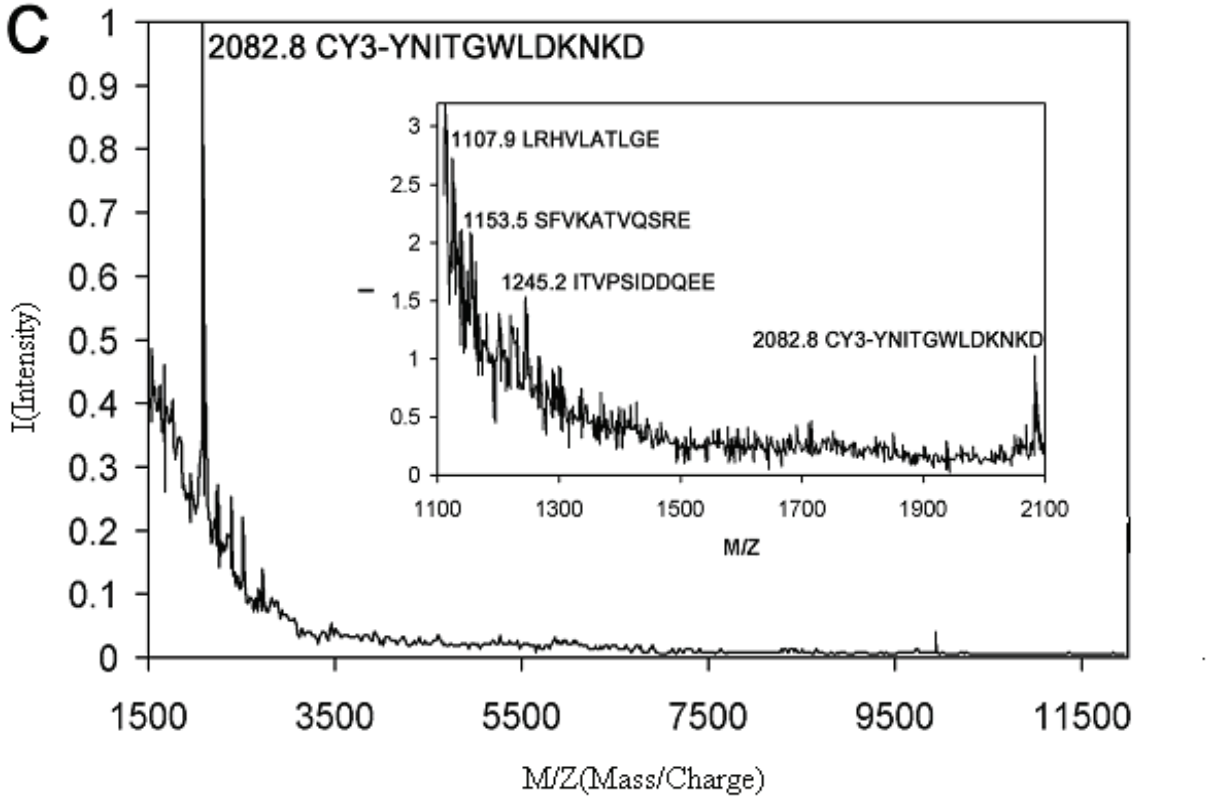

Fig 6.3C: The 2 kDa fluorescent fragment was extracted from the tricine gel and analyzed by mass spectroscopy. The determined mass of 2082.8 matched only a single predicted peptide-CY3 conjugate mass from S1, 590-601: YNITGWLDKNKD. This peptide is located near the C-terminal side of the 50 kDa subdomain agreeing with limited tryptic digestion analysis of the CY3-S1. No peptide masses corresponding to possible CY3-labeled were detected in the plausible mass range as identified in Table 1. A few unlabeled peptide signatures were detected as corresponding to fragments from the S1 sequence.

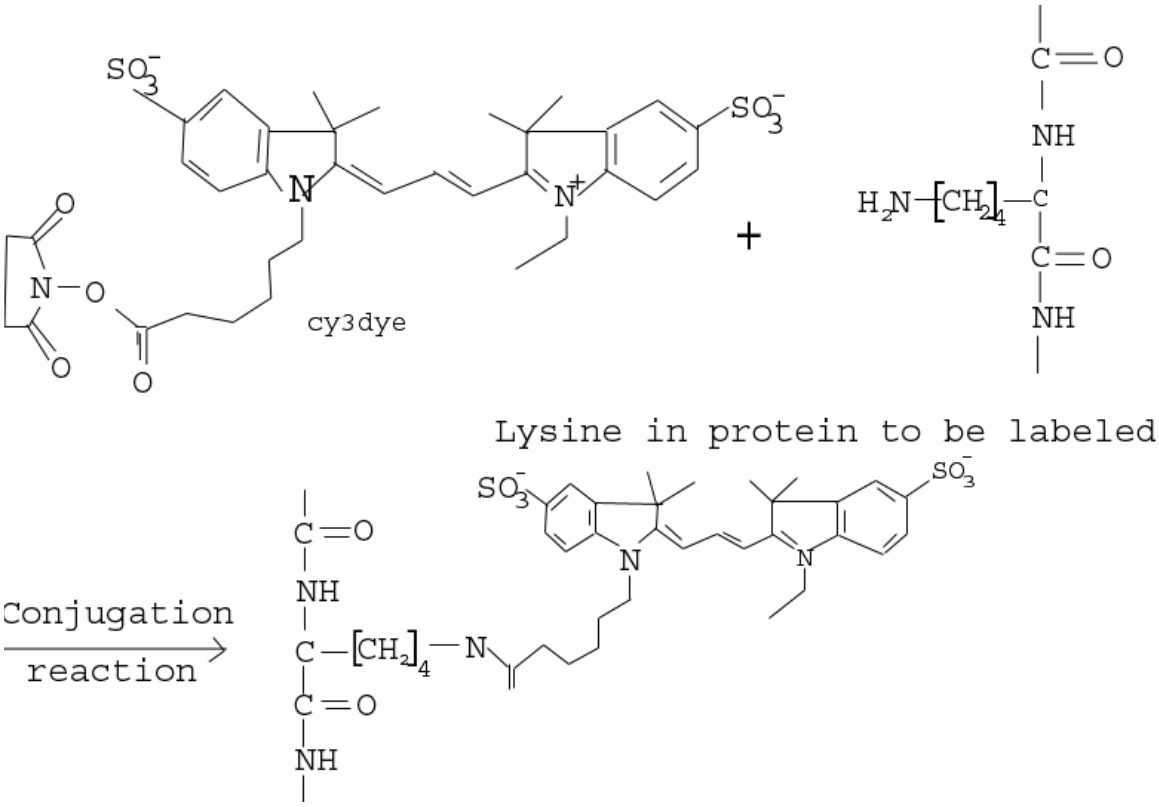

Fig. 6.3D: Chemical reaction labeling with cyanine dyes. The dye structure contains active ester that reacts with free primary amino groups of protein and will form a covalent link via an amide bond to the protein

Size with CY3	Size of the peptide	Sequence	Detection of labeled peptide	Detection of peptide only
1737	1123	ATSGKMQGTLE	No	No
1767	1153	SFVKATVQSRE	No	Yes
1821	1207	AFANAKTVRND	No	No
3126	2510	NSSRFGKFIRIHFGTTGKLASAD	No	No
1922	1307	KSRVTFQLKAE	No	No
2642	2027	RSYHIFYQIMSNKKPD	No	No
2309	1694	HYGNMKFKQKQRE	No	No
1895	1281	KAAYLQSLNSAD	No	No
2318	1703	LLKALCYPRVKVGNE	No	No
2339	1724	QQVYNAVGALAKAVYE	No	No
2538	1923	KMFLWMVTRINQQLD	No	No
2180	1565	TKQPRQYFIGVLD	No	No
2433	1818	KLQQFFNHHMFVLE	No	No
1182	567	YKKE	No[1]	No
1749	1135	KPMGIFSILE	No	No
1527	912	CMFPKATD	No	No
1744	1130	TSFKNKLYE	No	No
2950	2334	QHLGKSNNFQKPKPAKRKVE	No	No
2082	1467	YNITGWLDKNKD	Yes[2]	No
3234	2618	TVVGLYQKSAMKTLAFLFTGTAAAE	No	No

[1]This peptide was not detected as a distinct peak in the mass spectrum, even though a few unlabeled peptide masses were detected in this range suggesting that this hypothetical labeled peptide was not present. (See Fig. 6.3C.)

[2]This fragment also contains a possible KNKD fragment that apparently wasn't formed in significant amounts during the V8 protease digestion and would have had a CY3-labeled mass of 1,119 Da.

Near these residues on the surface of the atomic model of myosin is a patch of hydrophobic residues surrounded by basic amino acids. The plausibility of this location was tested by collisional quenching of the CY3-S1 with differently charged collisional quenchers (Fig. 6.4). Consistent with this hypothesis, the positively charged quencher Tl+ was less efficient than I- in quenching CY3 on S1 than CY3 free in solution indicating a positive electric potential in the vicinity of the CY3 dye on S1. In addition, the CY3-S1 was almost as accessible to collisional quenchers as the free CY3 indicating that the CY3 is located on the surface of S1.

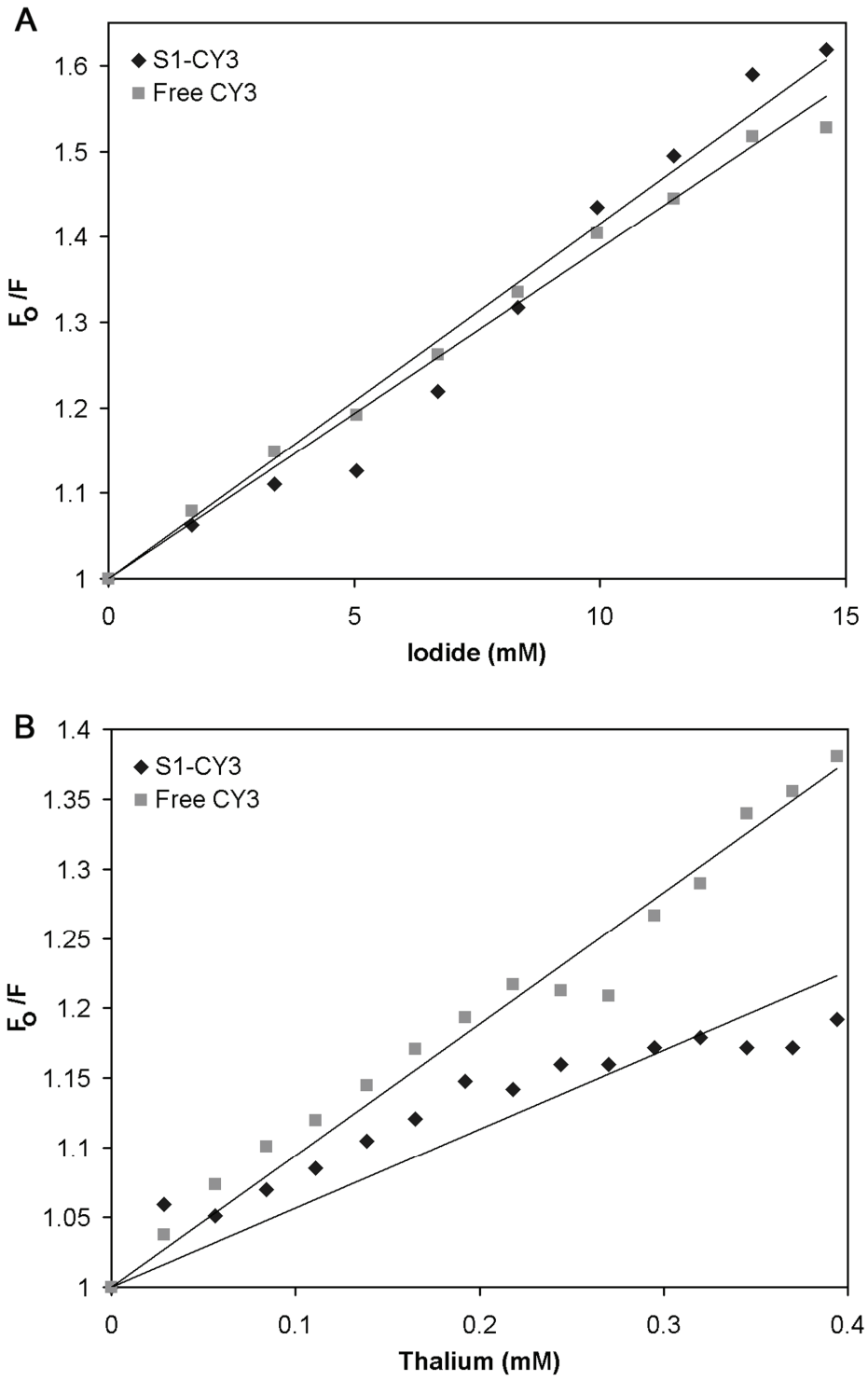

Fig. 6.4: Susceptibility of CY3-S1 fluorescence to collisional quenching. CY3-S1 and free CY3 had similar accessibilities to **(A)** I^- indicating that the CY3 is exposed on the surface of S1. The accessibility of CY3-S1 to **(B)** Tl^+ is 1.6-fold less than to free CY3 indicating positive electric potential in the vicinity of the CY3 on S1. Measurements were made in 0.1 M KCl, 2 mM $MgCl_2$, 10 mM imidazole, pH 7.0.

Given the location of the CY3 near the actin-binding cleft, it appeared likely that S1 binding to actin might perturb the properties of the CY3. This hypothesis was tested by measuring the fluorescence polarization of CY3-S1 (Fig. 6.5A). The data indicate that actin binding significantly reduced the fluorescence polarization of CY3 implying an increased mobility of the CY3. In contrast, nucleotide binding did not produce a detectable impact on the fluorescence polarization of CY3-S1. The fluorescence polarization of control free CY3 dye was also unaffected by the addition of nucleotides or actin.

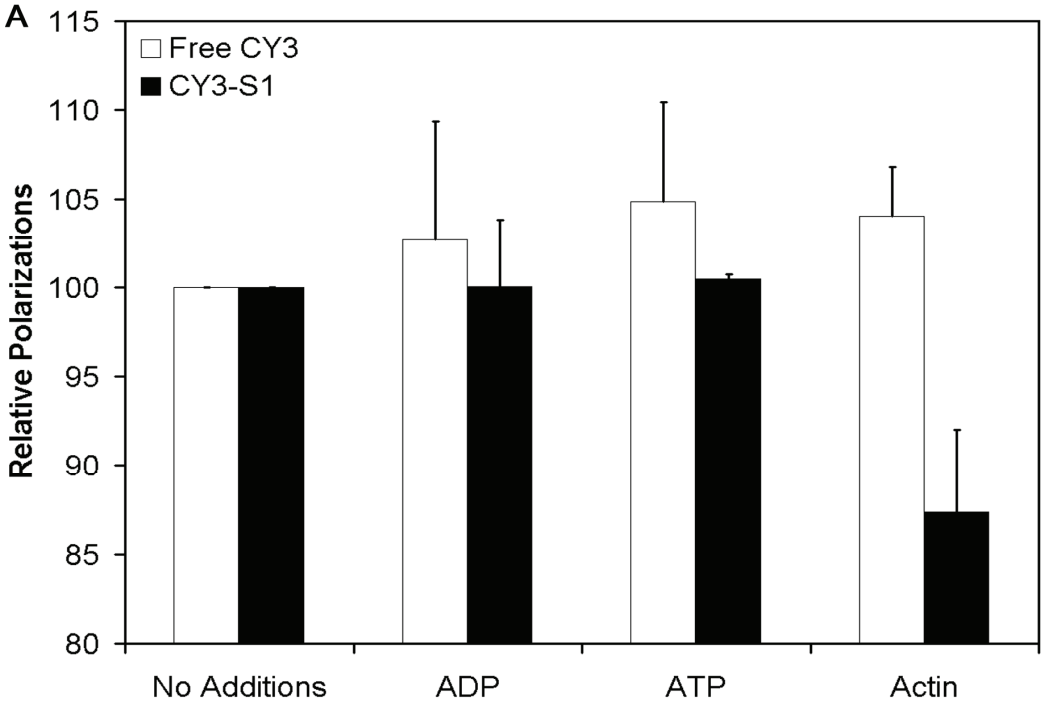

Fig. 6.5A: Actin binding to CY3-S1 detected by fluorescence polarization.The relative fluorescence polarizations values of 0.0005 mM CY3-S1 or free CY3 in the presence of no nucleotide or actin (No Additions), 0.001 mM ADP (ADP), 0.5 mM ATP (ATP), or 0.005 mM actin and ADP (Actin). Nucleotides did not cause a statistically significant change for either CY3-S1 or free CY3. Only the presence of actin induced a significant decrease in the fluorescence polarization. Since the fluorescence intensity at the magic angle did not appreciably change in the presence of nucleotides or actin, it is assumed that the fluorescence lifetime did not change indicating that actin caused an increase in mobility of the CY3 on S1. Measurements were made in 0.1 M KCl, 2 mM $MgCl_2$, 10 mM imidazole, pH 7.0.

To assess if the label might impact S1 function, the Mg^{2+} ATPase activities of S1 were assayed in the presence and absence of actin. These assays yielded activities of 0.06 s^{-1} and 2.9 s^{-1} for CY3-S1 and 0.06 s^{-1} and 2.7 s^{-1} for unlabeled S1 in the absence and presence of 125 μM actin, respectively. The unlabeled S1 had 18 μM Km, and 3.6 s^{-1} Vmax, but the Km and Vmax of CY3-S1 could not be precisely determined due to the apparently low affinity of actin for CY3-S1. Still the activity was not decreased by the CY3 labeling as suggested by the double reciprocal plot analysis (Fig. 6.5B). Thus, the labeling did not substantially reduce the Vmax but did increase the Km for actin consistent with an impact at the actomyosin interface.

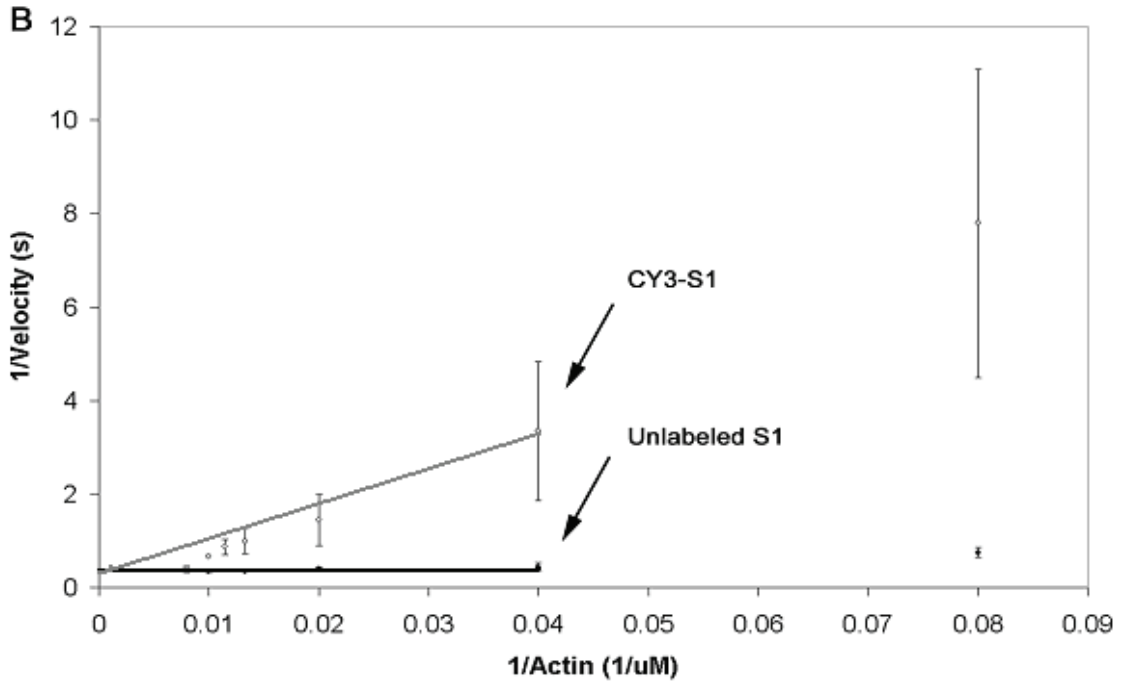

Fig 6.5B: Double reciprocal plot analysis of actin-activated ATPase activities of CY3-S1 and unlabeled S1 demonstrate reduced affinity of actin for CY3-S1 relative to unlabeled S1. The 3.6 s^{-1} V_{max} of unlabeled S1 was less than or equal to the V_{max} of CY3-S1. The 18 μM K_m of unlabeled S1 compared to a K_m of greater than 100 μM for CY3-S1. The bars represent the deviation of values obtained for seven samples. The deviation is greater at lower concentration so a standard fit was not applied for lower concentrations. Actin activated ATPases were assayed in 2 mM $MgCl_2$, 10 mM imidazole, pH 7.0 and 25^0C.

CHAPTER 7

COMPUTATIONAL RESULTS

Procedure for Opening /Closing 50 kDa Cleft

Atomic model of native myosin S1 is energy minimized. To the native energy minimized structure, eleven pairs of harmonic constraints are applied evenly between the upper and lower subdomains around the cleft. Each distance constraint is incremented/ decremented by 0.1 nm followed by energy minimization. This process is continued for 12 cycles. Molecular dynamics simulations are performed for 100 ps to the final energy minimized structure. The final energy minimized structure of the opened/closed 50 kDa cleft is compared to the energy minimized structure of native myosin S1. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

Simulations of the 50 kDa Cleft

Since the strut is the primary sequence spanning the 50 kDa cleft, it is hypothesized that movements of this cleft could perturb the strut sequence. In order to understand how the structural changes in myosin S1 and the strut are affected by opening and closing of the large 50 kDa cleft, we used molecular mechanics and dynamics simulations where harmonic constraints were applied around the cleft so that the cleft is forced to either close or open. Flexible surface loops were modelled into the crystal structure of the motor domain portion of skeletal myosin S1 using local conformational searching techniques. The 50 kDa cleft required somewhat more force (work) to close than to open. The force required during the initial stages for opening and closing the 50 kDa cleft is nearly same suggesting that the cleft is equally predisposed to movement in either direction during the

initial stages (Fig. 7.3), but during the later stages the force required to close the 50 kDa cleft is more. Superimpositions of the constrained cleft structures with native structures gives root mean square deviation(RMSD) which revealed substantial changes at the alpha-helix leading to switch 2, loop 1, loop 2 and the sequence leading up to the reactive cysteines (Fig. 7.1 A, 7.2 A). Closing the 50 kDa cleft had a greater impact on the strut sequence than opening the cleft. While impacted by the cleft movements, the portion of the strut sequence where CY3 is attached demonstrated a lower magnitude of induced change.

Opening of 50 kDa cleft has more effect on myosin S1 than closing the 50 kDa cleft. The magnitude and number of changes is considerably more. Molecular dynamics simulations indicate that there are more changes in the lower 50 kDa domain than the upper 50 kDa domain (Fig 7.1 B, 7.2B). The alpha-helix leading to the lever arm takes a swing of about 6 Å when the cleft is opened (Fig. 7.4), while there is hardly any change in the position of the lever arm when the cleft is closed. The lever arm does not move in the initial stages of cleft opening, it moves only after the cleft is opened by more than 2 nm. Although interactions of the SH3 domain with the essential light chain might potentially have some impact on the lever arm. Somewhat more of an effect is apparent on the structures surrounding the nucleotide cleft.

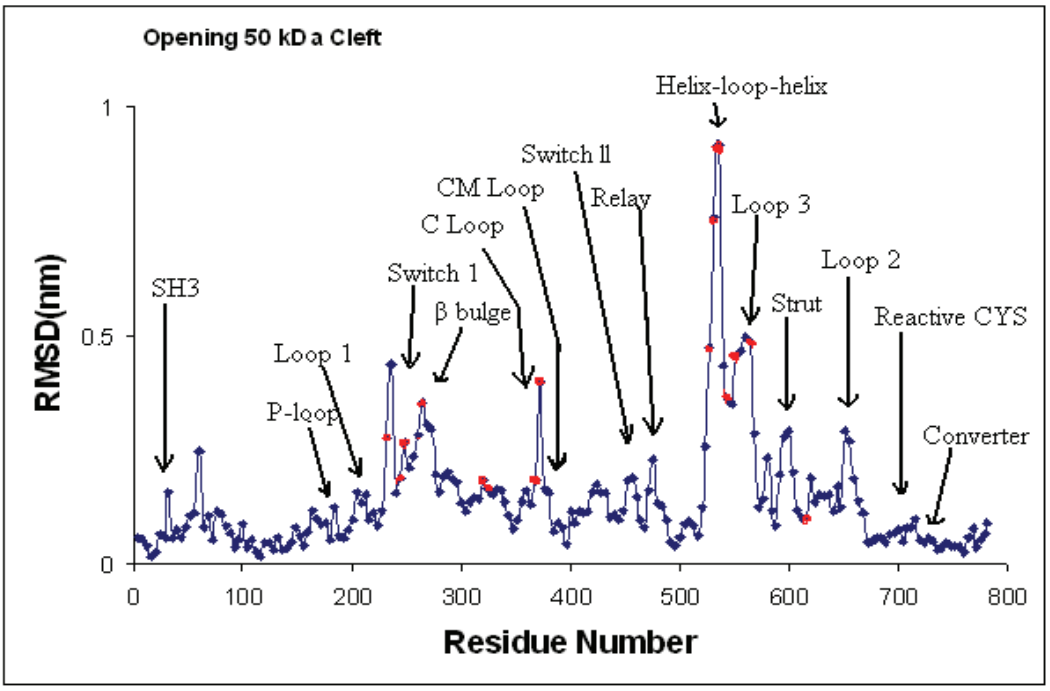

Fig. 7.1 A: Structural changes associated with opening of the 50 kDa cleft. The endpoints of energy minimised structure is compared to the native energy minimised structure of myosin S1 and the magnitude of changes is determined from root mean square deviation(RMSD).

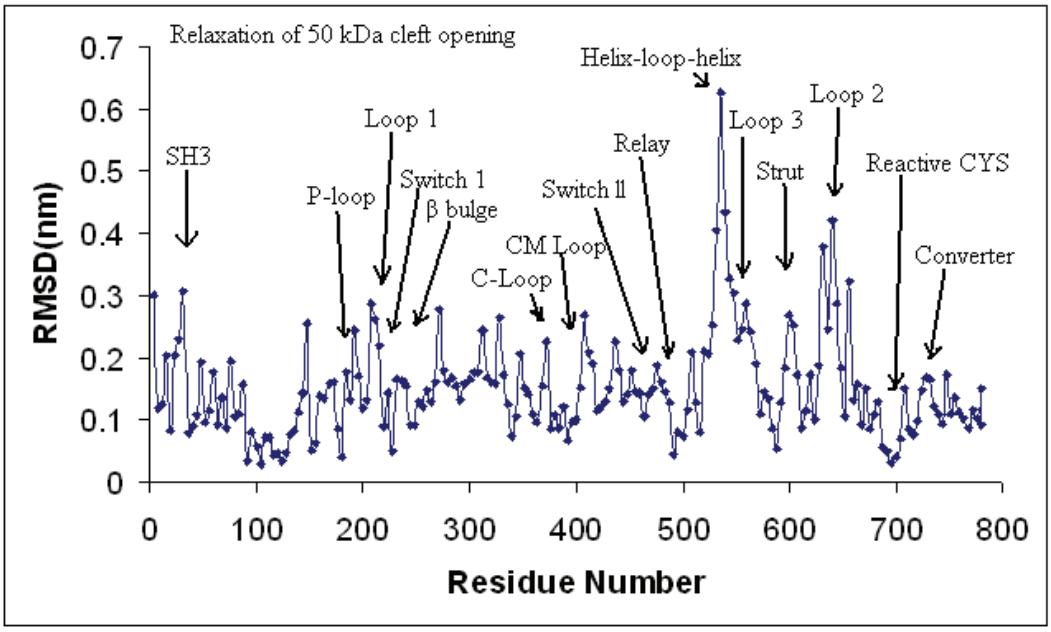

Fig. 7.1 B: Structural changes after simulating the endpoints of energy minimized structure of open 50 kDa cleft for 100 ps without constraints during relaxation.
The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

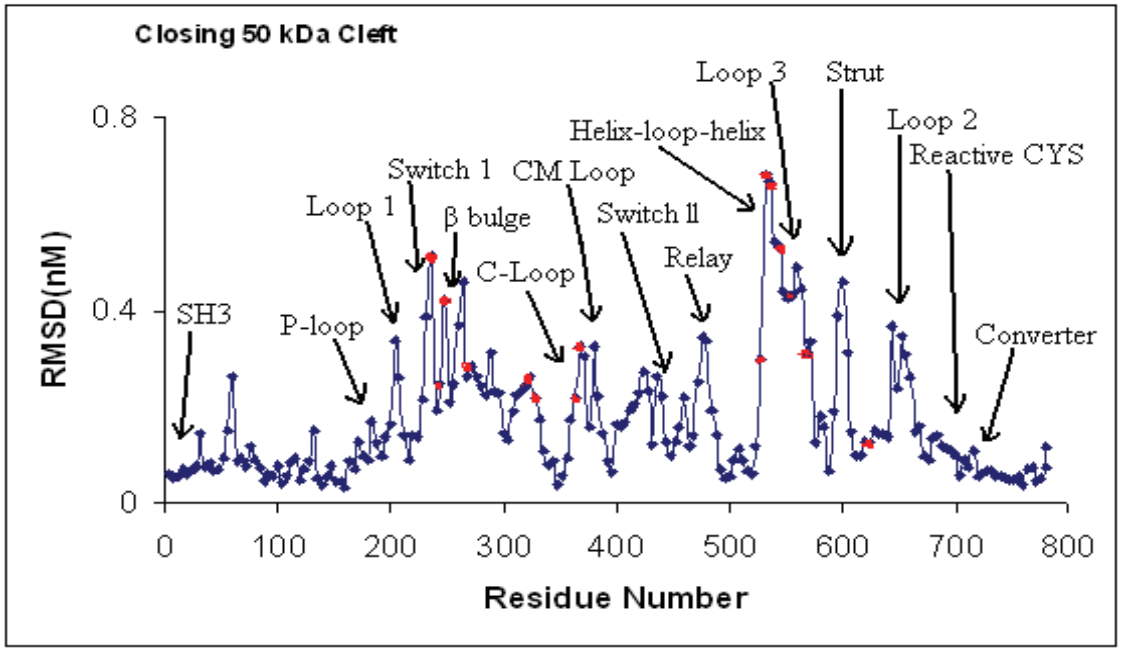

Fig. 7.2 A: Structural changes associated with closing of the 50 kDa cleft. The endpoints of energy minimized structure is compared to the native energy minimized structure of myosin S1.

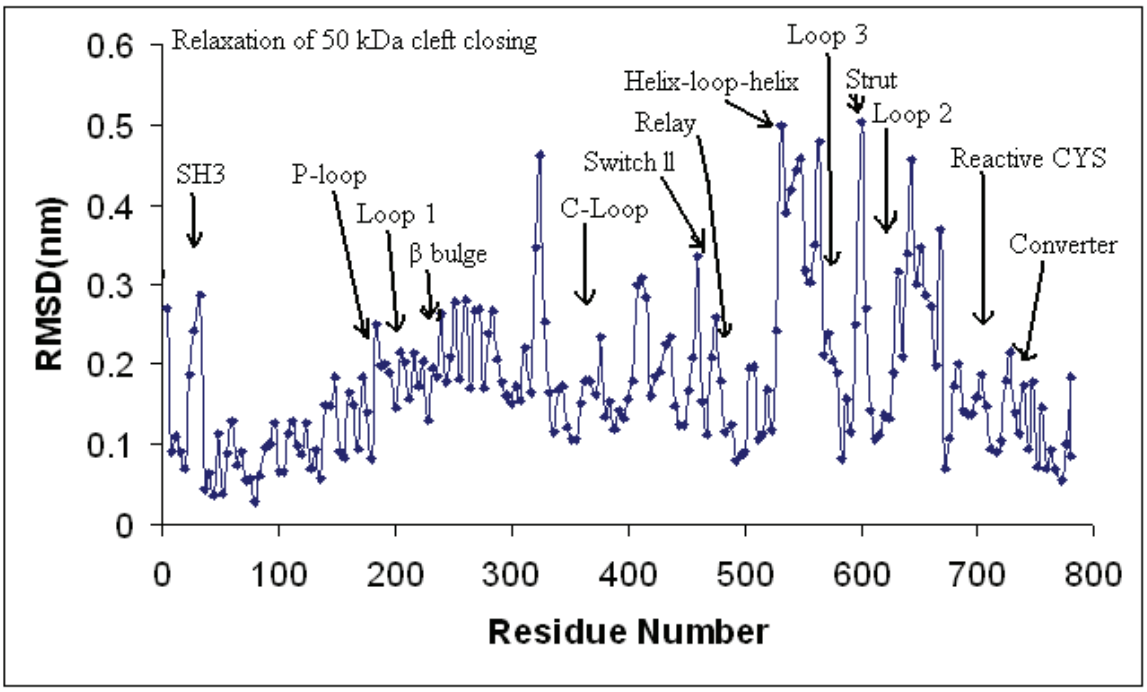

Fig. 7.2 B: Structural changes after simulating the endpoints of energy minimized structure of closed 50 kDa cleft for 100 ps without constraints during relaxation. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

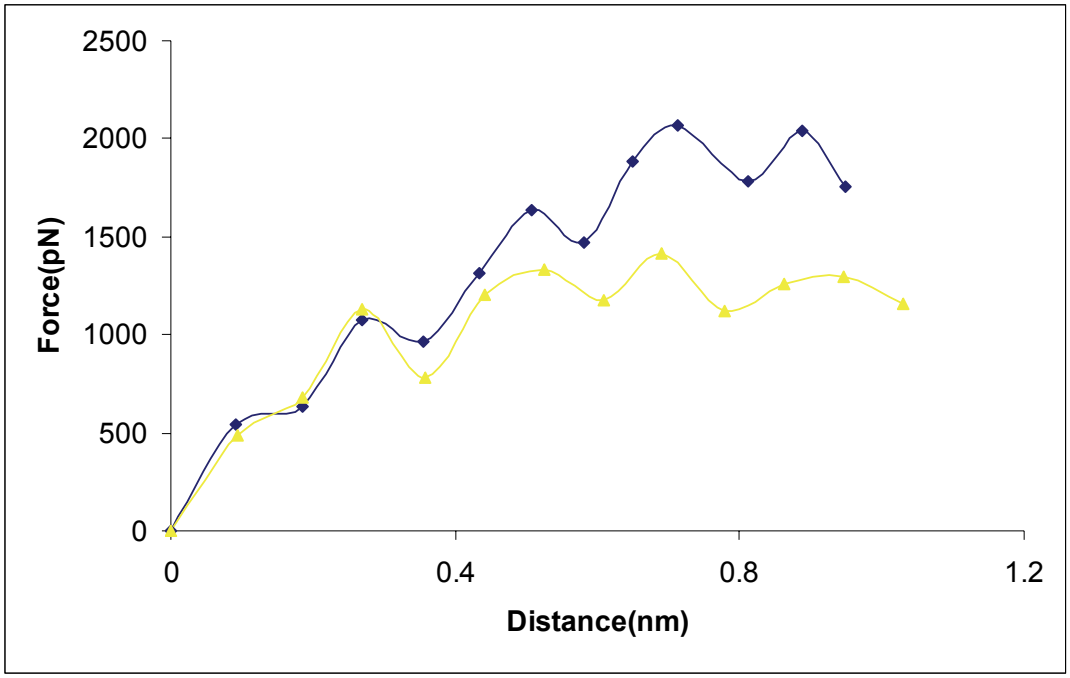

Fig. 7.3: Simulated closing and opening of the 50 kDa cleft. Force constraints with a spring constant of 16.7 N/m were used to close and open the 50 kDa cleft by changing the target distance at 0.1 nm increments followed by energy minimization. A simulated force–distance curve was calculated based on the difference between the actual distance and the target distance after each energy minimization using the spring constant.

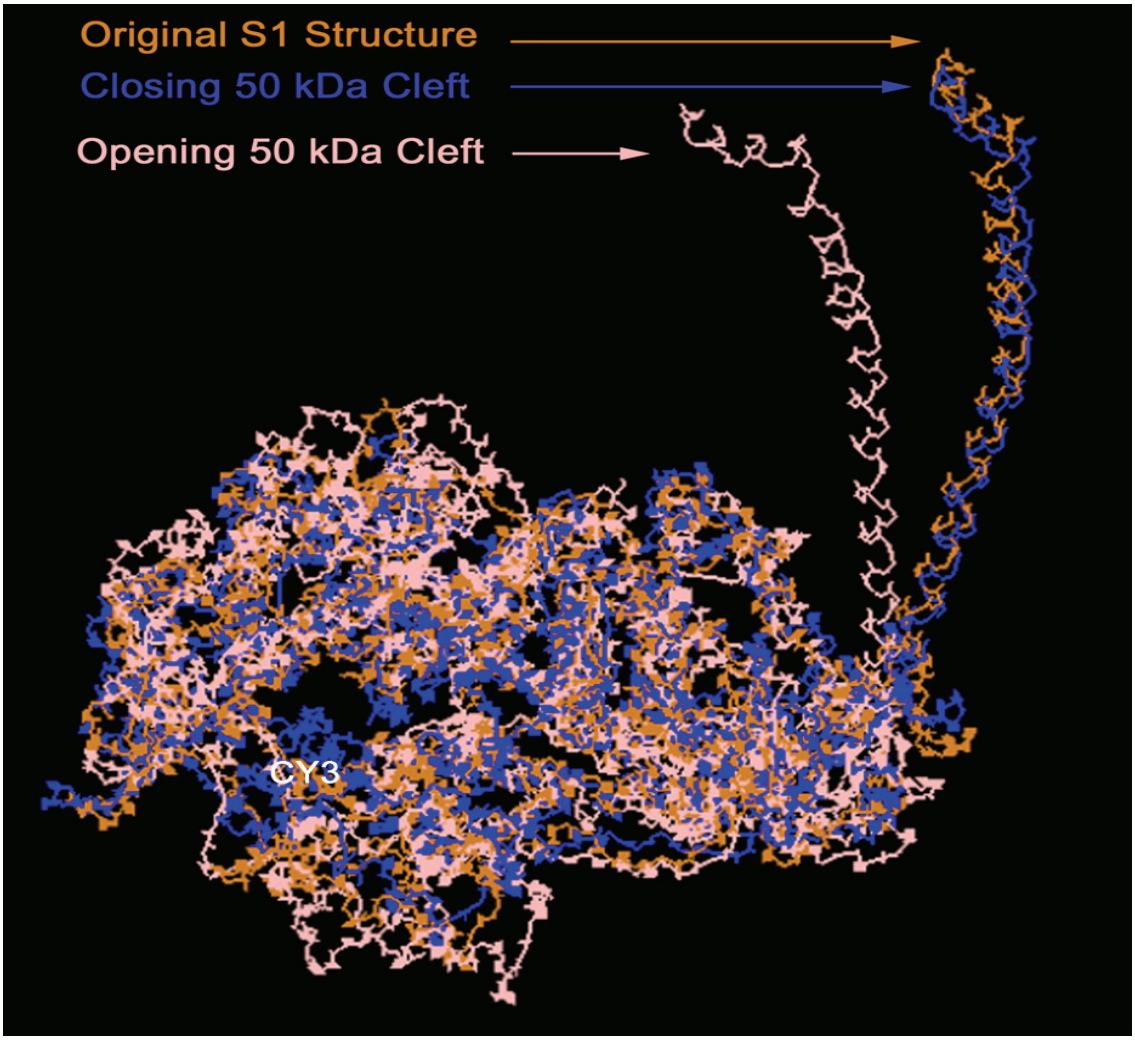

Fig.7.4: Effects of 50 kDa cleft closing and opening on the extrapolated lever arm position and the position of the CY3 label. The alpha-helix of the lever arm was reattached post-simulation to illustrate the impact of closing and opening the original 50 kDa cleft. Each model is superimposed with the original atomic model. The induced motions are relatively small, but the amplified effect of opening the cleft on the lever arm position is substantial [29,40,103,111]. More modest changes are apparent in the position of the CY3 labeling site on the strut (indicated in white text.)

Procedure for SH1-SH2 Crosslinking.

Atomic model of native myosin S1 is energy minimized. To the native energy minimized structure, distance constraint is applied to CYS 697 and CYS 707. The two reactive cysteines are separated by a distance of 2 nm. The two cysteines are pulled closer by 0.1 nm during each energy minimization cycle. Force constraints with a spring

constant of 16.7 N/m were used to crosslink the two cysteines by changing the target distance at 0.1 nm decrements followed by energy minimization. This process is continued till the two cysteines get to a distance of 0.3 nm. This entire process takes 34 energy minimization cycles. Molecular dynamics simulations are performed for 100 ps to the final energy minimized structure with the two cysteines at 0.3 nm with constraints attached to the two cysteines during simulation and also without any constraints attached to the two cysteines. The final energy minimized structure of the crosslinked cysteines is compared to the energy minimized structure of native myosin S1. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

Simulations near the Converter Region

Crosslinking SH1-SH2 with the nucleotide binding site closed does not show any changes in the structure (Fig 7.5 A). The CY3 labeling site was unperturbed by crosslinking C697 and C707. Crosslinking the reactive cysteines C697 and C707 have previously been shown experimentally to cause reduced actin binding affinity and trapping of the nucleotide in the active site. A force constraint was applied between these residues to computationally crosslink them in the atomic model. The residues are located on the opposite ends of a short -helix in the catalytic domain of the myosin head and are separated from one another by ~20 Å. The mechanism by which this helix is involved in the force generation cycle is still unknown.

Crosslinking the reactive cysteines C697 and C707 induced movements that correlated with modest changes throughout the motor domain particularly in the normally flexible regions identified by the dynamics simulations (Fig 7.5 B, 7.5 C). The structure

does not get back to original RMSD during relaxation, which suggests that the crosslinking of cysteines causes a change in flexibility. This is also seen during many cysteine crosslinking experiments like loss of ATP ase activity inactivation of motility and trapping of nucleotide [105]. The strut region seems to be the most rigid region in the motor domain as it does not get perturbed and has one of the lowest RMSD. It has previously been shown that the strut is very rigid as it acts to keep the upper and lower 50 kDa domain in disposition.

Crosslinking caused little change in the helix leading to the lever arm. If conformational changes from crosslinking were directly involved in lever arm movement, then the computational crosslinking should have affected this helix. Indirect positioning of the lever arm through light chain interactions with the converter region and the SH3 domain seem more consistent with the simulations. It has been suggested that the light chain binding domain is required to see the expected conformational change at the fulcrum point.

The force-distance plot shows that the crosslinking of reactive cysteines requires substantial force as the structure has to undergo some major barriers like the unraveling of the alpha-helix that separates the two residues and so the force rises during few stages and then drops (Fig 7.6).

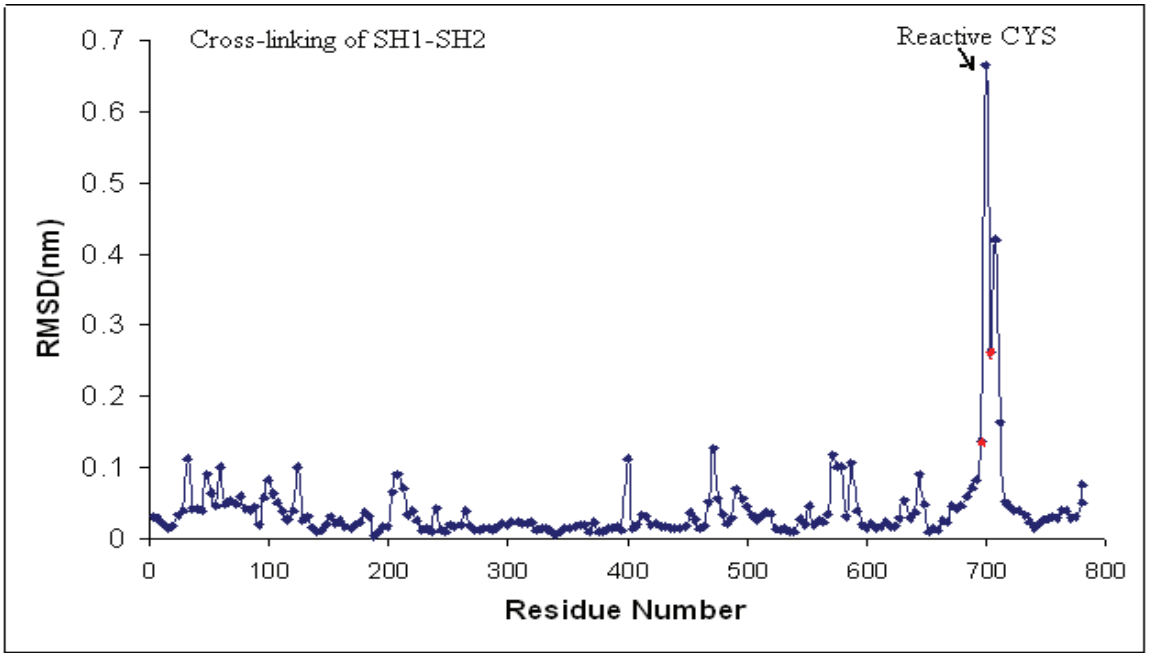

Fig. 7.5 A: Structural changes associated with crosslinking of the reactive cysteines SH1 and SH2 with nucleotide binding site closed. The endpoints of cysteine crosslinked energy minimised structure is compared to the native energy minimised structure of myosin S1.

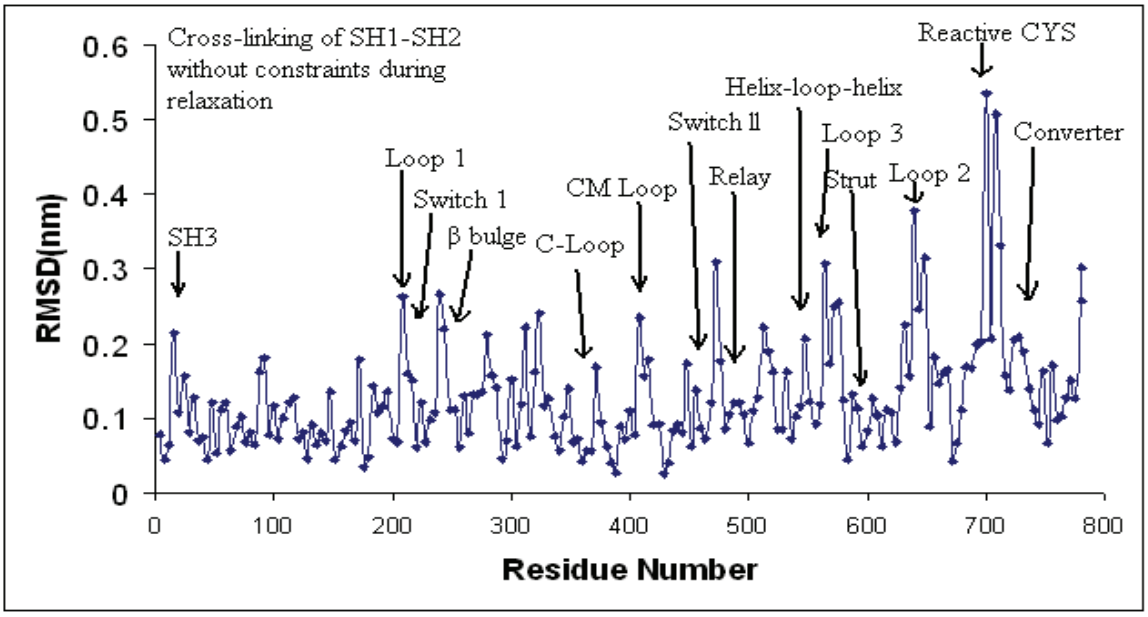

Fig. 7.5B: Structural changes after simulating the endpoints of energy minimized structure of crosslinked cysteines with nucleotide binding site closed for 100 ps without constraints during relaxation. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

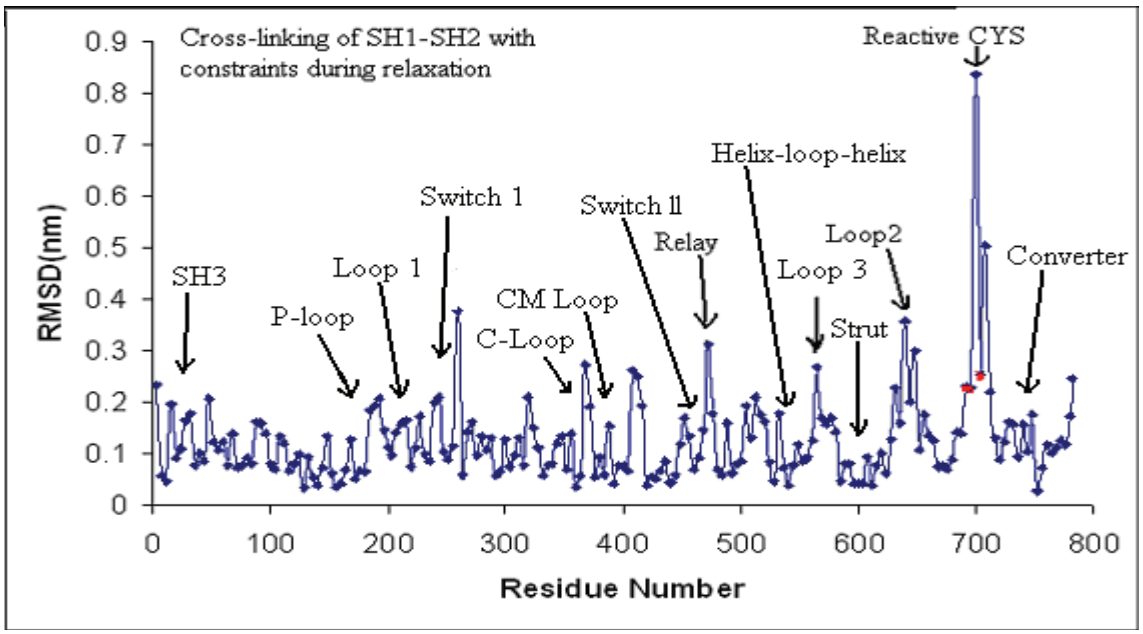

Fig. 7.5 C: Structural changes after simulating the endpoints of energy minimized structure of crosslinked cysteines with nucleotide binding site closed for 100 ps with constraints at CYS 697 and CYS 707 during relaxation. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

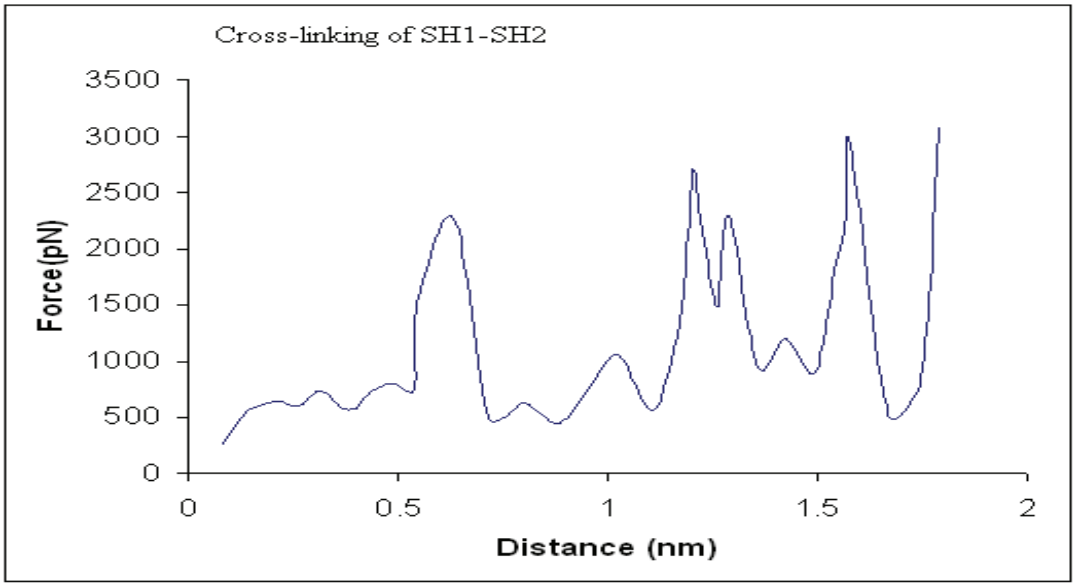

Fig. 7.6: Force-distance curve of SH1-SH2 crosslinking with nucleotide binding site closed. Force constraints with a spring constant of 16.7 N/m were used to crosslink the reactive cysteines by changing a distance of 0.1nm decrements followed by energy minimization. The force-distance curve was determined by calculating the difference between the actual distance and the target distance after each energy minimization using the spring constant.

Procedure for Opening Nucleotide Binding Site.

Atomic model of native myosin S1 is energy minimized. To the native energy minimized structure, distance constraint is applied to SER 180 and SER 320. The two serines are separated by a distance of 1.5 nm and are a part of the phosphate binding site. The two serines are pulled apart by 0.1 nm during each energy minimization cycle. Force constraints with a spring constant of 16.7 N/m were used to open the nucleotide binding site by incrementing the target distance at 0.1 nm followed by energy minimization. This process is continued till the two serines are separated apart by a distance of 2.5 nm from 1.5 nm. This entire process takes around 17 energy minimization cycles. Molecular dynamics simulations are done for100 ps to the final structure with the two serines at 2.5 nm with constraints attached to the two serines during simulation and also without any constraints attached to the two serines. The final energy minimized structure of the nucleotide opened structure is compared to the energy minimized structure of native myosin S1. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

Simulations of the Nucleotide Binding Site

There are no major structural changes in S1 when nucleotide binding site is opened (Fig. 7.7 A). The CY3 labelling site was strikingly unperturbed by the forces applied to the nucleotide cleft which is consistent with the lack of observed effects of nucleotide binding on the fluorescence polarization measurements (Fig. 6.5). The single force constraint between residues I182 and S325 induced nucleotide cleft movements that correlated with modest changes throughout the motor domain particularly in the normally flexible regions identified by the dynamics simulations (Fig 7.7B, 7.7 C). The highest

RMSD was near the cardiomyopathy loop, when the dynamics simulations were performed for 100 ps with constraint between residue I182 and S325, but it is not obvious that the change is from cause and effect.

With the constraint released from the residue I182 and S325 and dynamics simulations performed for 100 ps there were again modest changes throughout the motor domain particularly in the normally flexible regions. Switch 1, which is not a flexible joint shows the highest RMSD. This is because the nucleotide-binding pocket is located at the interface between the N-terminal and 50-kDa upper subdomains, opening of the nucleotide binding pocket causes the rotation of the 50-kDa upper subdomain and that the center of rotation is located in the switch I region which was previously observed experimentally in scallop myosin.

In the presence and absence of constraint between residues I182 and S325 during dynamics simulations the strut does not get perturbed much but shows more RMSD when compared to crosslinking to the reactive cysteines. It seems that the strut is more prone to get disturbed from the opening of nucleotide cleft than crosslinking of reactive cysteines.

Opening of nucleotide site caused little change in the helix leading to the lever arm. If conformational changes from nucleotide site were directly involved in lever arm movement, then the computational opening of the nucleotide site should have affected this helix. Indirect positioning of the lever arm through light chain interactions with the converter region and the SH3 domain seem more consistent with the simulations.

The force (work) required to open the nucleotide site is very less, compared to crosslinking the reactive cysteines (Fig. 7.8). This is expected as it is a normal process during muscle contraction to release ATP when myosin S1 binds to actin.

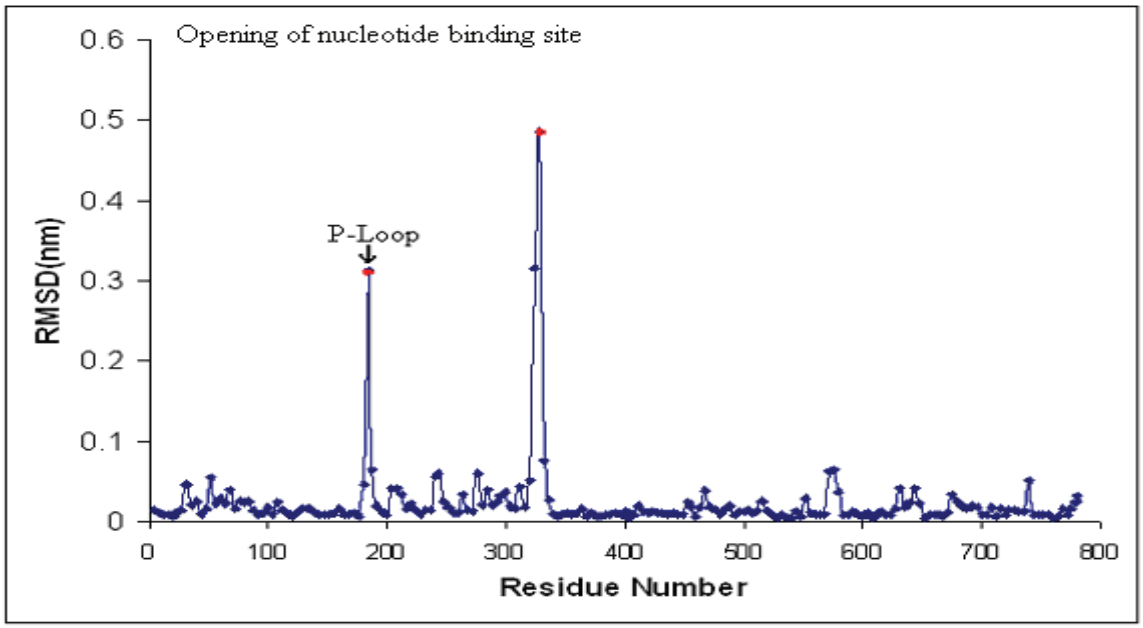

Fig. 7.7 A: Structural changes associated with opening of nucleotide site. The endpoints of energy minimised nucleotide opened site structure is compared to the native energy minimised structure of myosin S1.

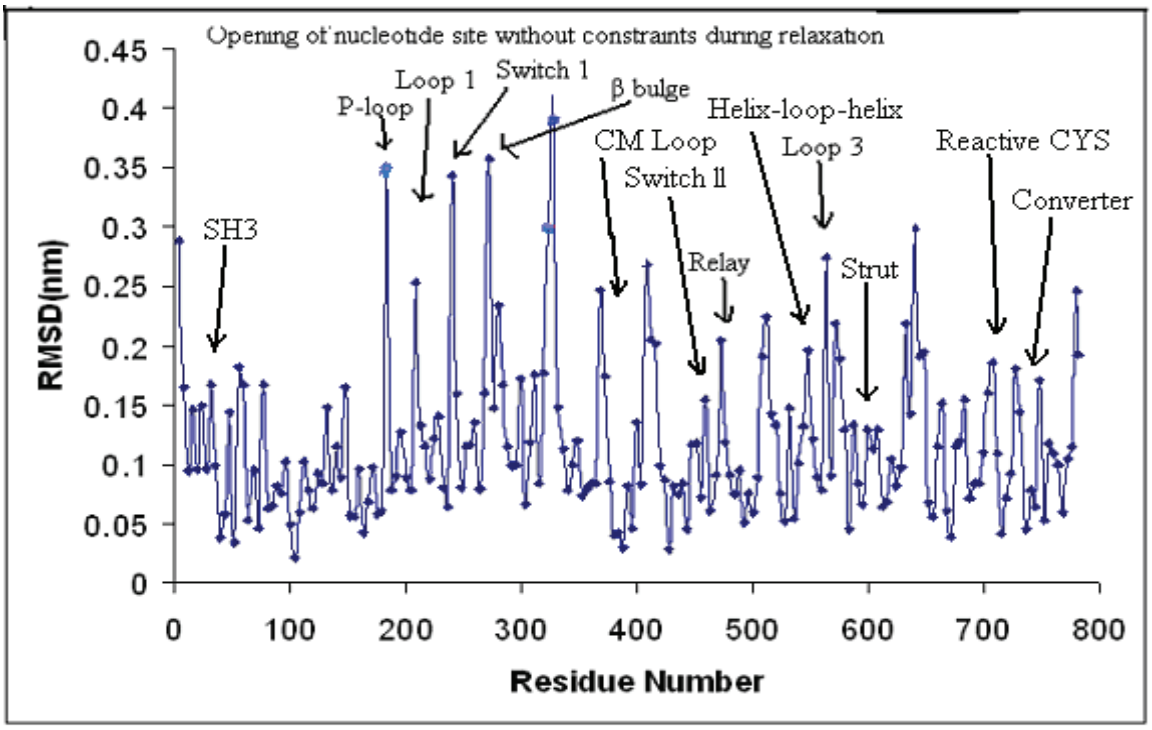

Fig. 7.7 B: Structural changes after simulating the endpoints of energy minimized structure of nucleotide opening site for 100 ps without constraints during relaxation. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

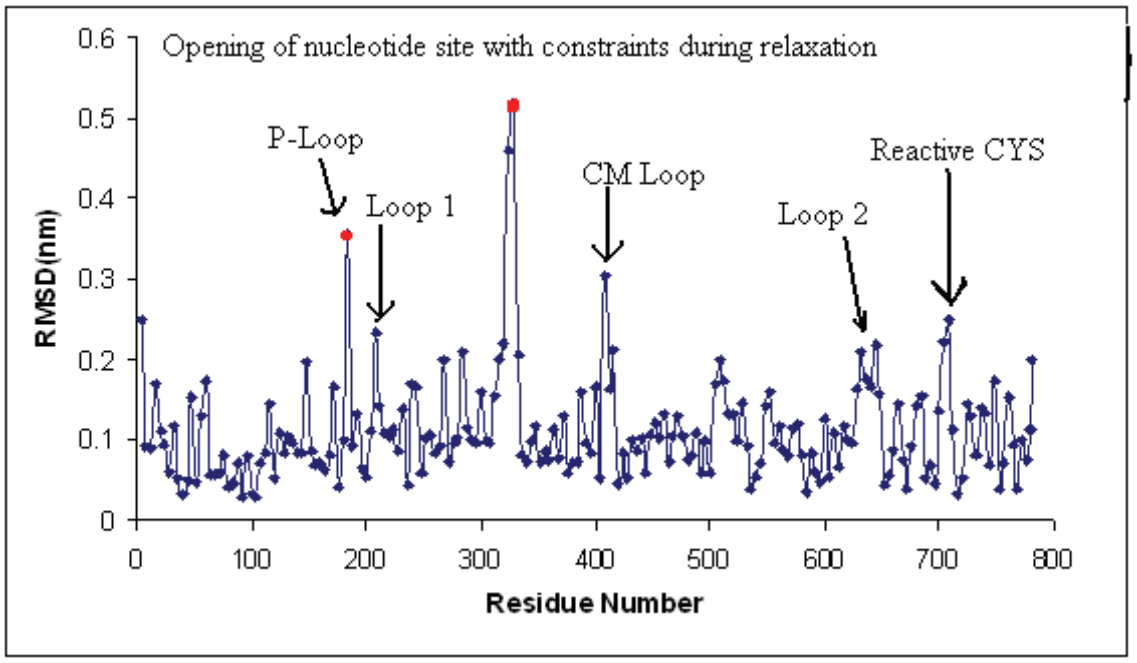

Fig. 7.7 C: Structural changes after simulating the endpoints of energy minimized structure of nucleotide opening site for 100 ps with constraints during relaxation. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 100 ps of molecular dynamics simulations.

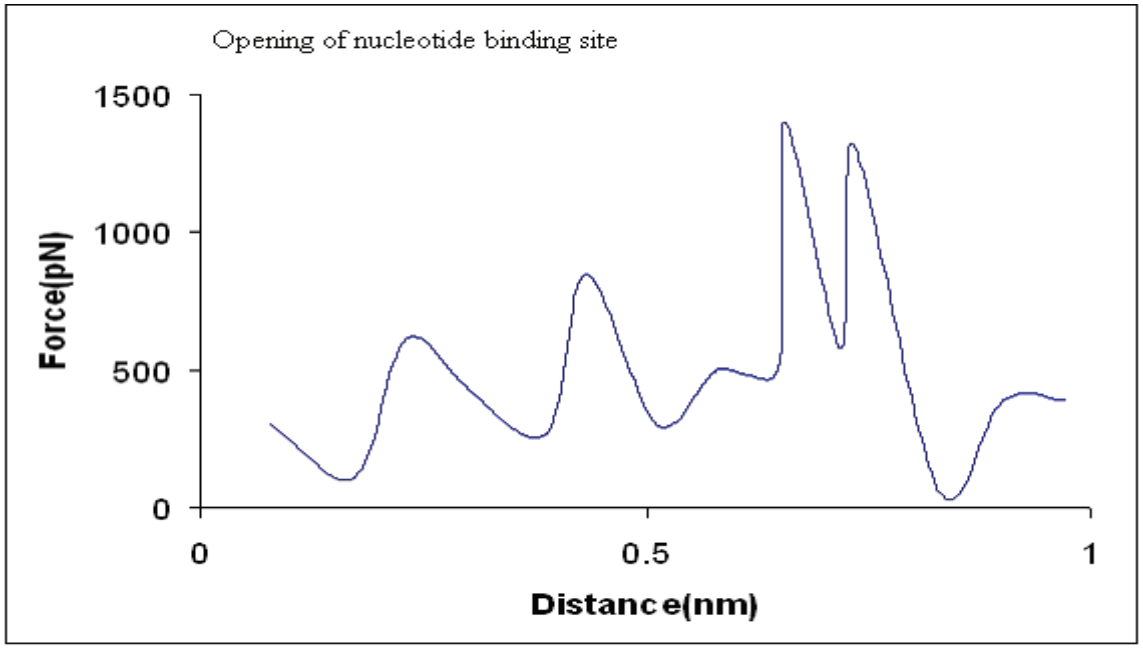

Fig. 7.8: Force-distance curve of nucleotide binding site. Force constraints with a spring constant of 16.7 N/m were used to open the nucleotide site SER 180 and SER 324 by changing a distance of 0.1nm increments followed by energy minimization.

Procedure for Crosslinking SH1 SH2 with Nucleotide Binding Site Open

The structure obtained from 100 ps of molecular dynamics simulations with constraints attached to the two serines (nucleotide binding site) is taken as the starting structure for crosslinking SH1-SH2. To this structure the reactive cysteines are crosslinked. Distance constraints are applied to CYS 697 and CYS 707 which are separated by a distance of ~2 nm. The two cysteines are pulled closer by 0.1nm during each energy minimization cycle and the two serines are also constrained and kept at a distance of 2.5 nm so that the nucleotide binding site is not closed. This process is continued till the two cysteines get to a distance of 0.3 nm. This entire process takes 31 energy minimization cycles. Molecular dynamics simulations are done for 100 ps to the final structure with the two cysteines at 0.3 nm and the two serines at 2.5 nm with constraints attached to the two cysteines and the two serines during simulation. Molecular dynamics simulations is also done without any constraints to the two cysteines and the two serines. The final energy minimized structure is compared to the energy minimized structure of native myosin S1 obtained after 100 ps of simulation. The structure obtained from molecular dynamics simulations with and without constraints is compared to the simulated structure of native myosin S1 obtained from 200 ps of molecular dynamics simulations.

Simulations near Converter Region and Nucleotide Binding Site

When the reactive cysteines are pulled into close proximity after opening nucleotide cleft, alpha-helix that separates the two residues unravels. A number of regions throughout the structure show modest sensitivity to the perturbation even though they are nanometers away from the perturbed site (Fig. 7.9 A). Such peaks in the plot of

RMSD correspond to regions of known significance such as the loop 1, loop 2, the converter region, and the helix between the CM loop, switch 1 and switch II. The CY3 labeling site which did not show any perturbation during crosslinking the reactive cysteines or opening of the nucleotide cleft shows significant RMSD when forces are applied to the nucleotide cleft in combination with forces applied to the converter region.

When the structure is entered into dynamics simulations for 100 ps in presence and absence of constraints none of the key regions relax back to typical RMSD levels but regions near to the site where forces are applied, switch 1 which is closer to the nucleotide binding site and loop 2 which is closer to the reactive cysteines increase in RMSD when constraints are removed during dynamics simulations (Fig 7.9 B, 7.9 C). This relaxation suggests that the strain on the regions where forces were applied is transmitted to the surrounding regions as the RMSD on the regions where constraints were placed decreases.

Crosslinking the reactive cysteines with nucleotide binding site open requires more work than with the nucleotide binding site closed (Fig. 7.10). Therefore, with nucleotide binding site closed, the crosslinking is accelerated and the helix undergoes some conformational changes in order for SH1 and SH2 to come close to each other. This could be indicative of helix melting or increased flexibility with nucleotide binding site closed. Previous crosslinking experiments also showed that nucleotides shift the equilibria among conformational states of the helix [112].

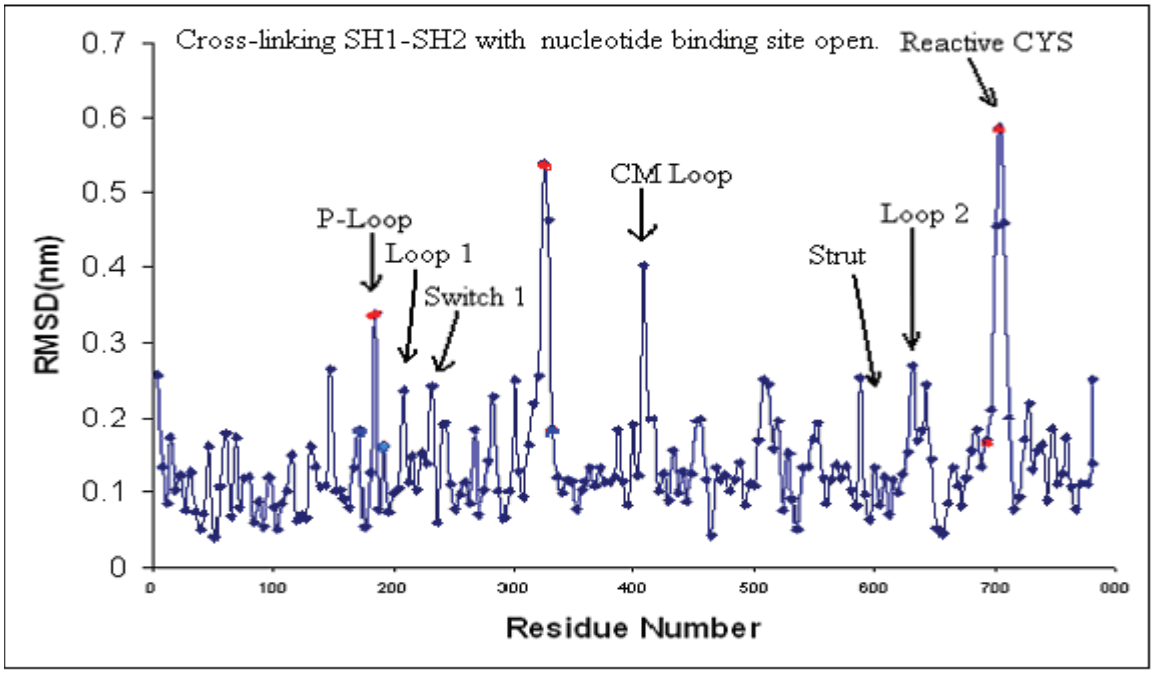

Fig. 7.9 A: Structural changes associated with crosslinking of the reactive cysteines SH1 and SH2 with nucleotide binding site open. The endpoints of cysteine crosslinked energy minimized structure after opening nucleotide binding site is compared to the energy minimized structure of native myosin S1 after 100 ps of molecular dynamics simulations.

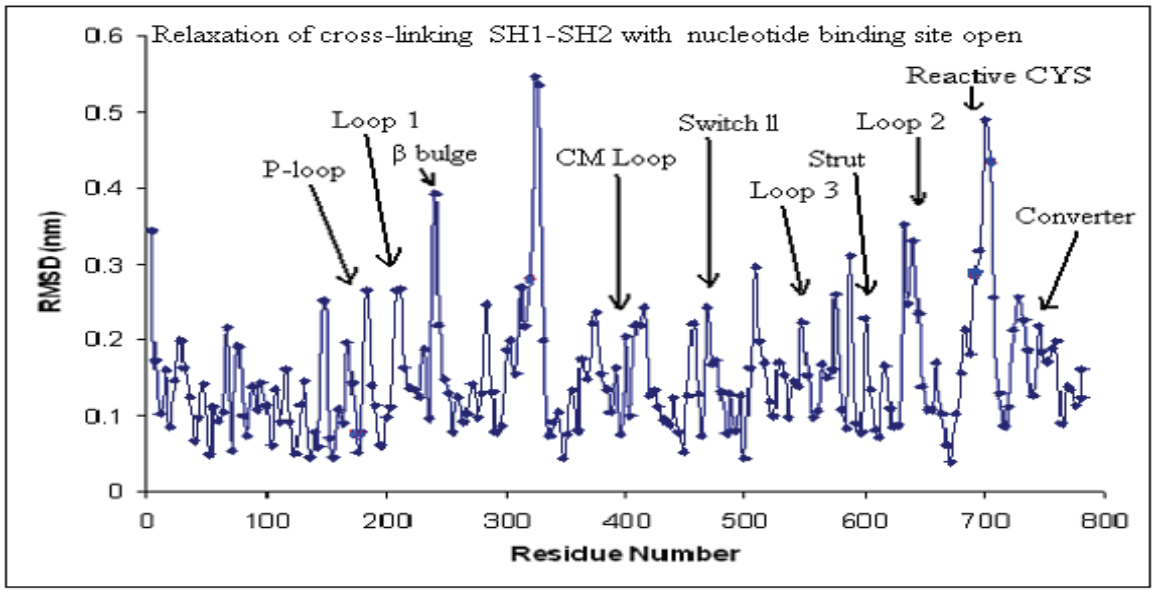

Fig. 7.9 B: Structural changes after simulating the endpoints of energy minimized structure of crosslinking of the reactive cysteines SH1 and SH2 after opening the nucleotide binding site with constraints during relaxation. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 200 ps of molecular dynamics simulations

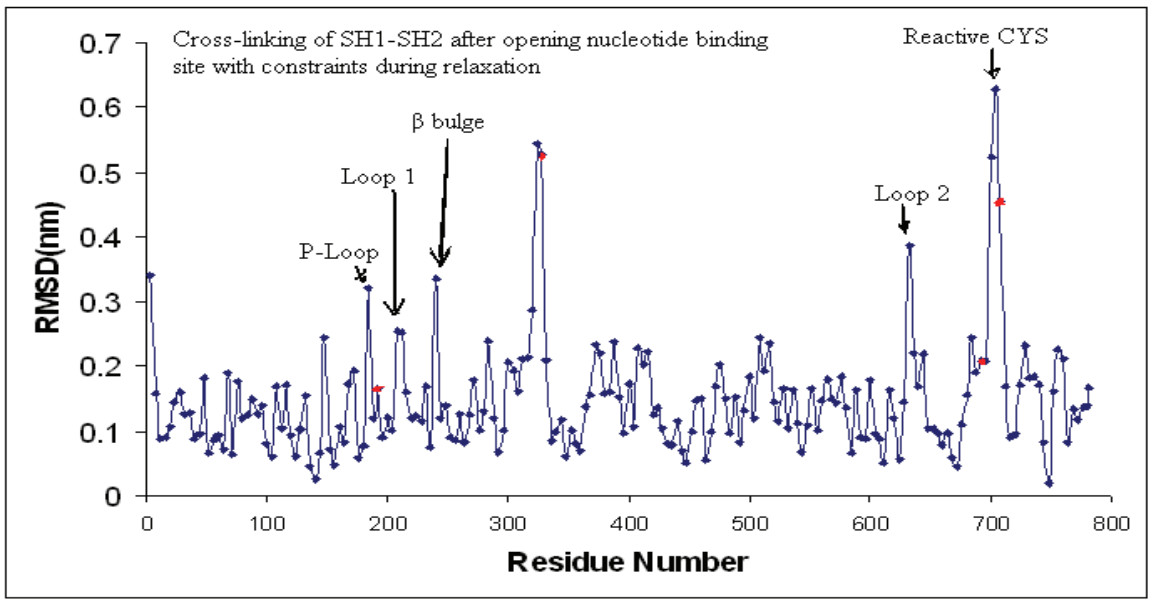

Fig. 7.9 C: Structural changes after simulating the endpoints of energy minimized structure of crosslinking of the reactive cysteines SH1 and SH2 after opening the nucleotide binding site with constraints at SER 180 and SER 324 and at CYS 697 and CYS 707 during relaxation. The structure is compared to the molecular dynamic simulated structure of native myosin S1 obtained after 200 ps of molecular dynamics simulations.

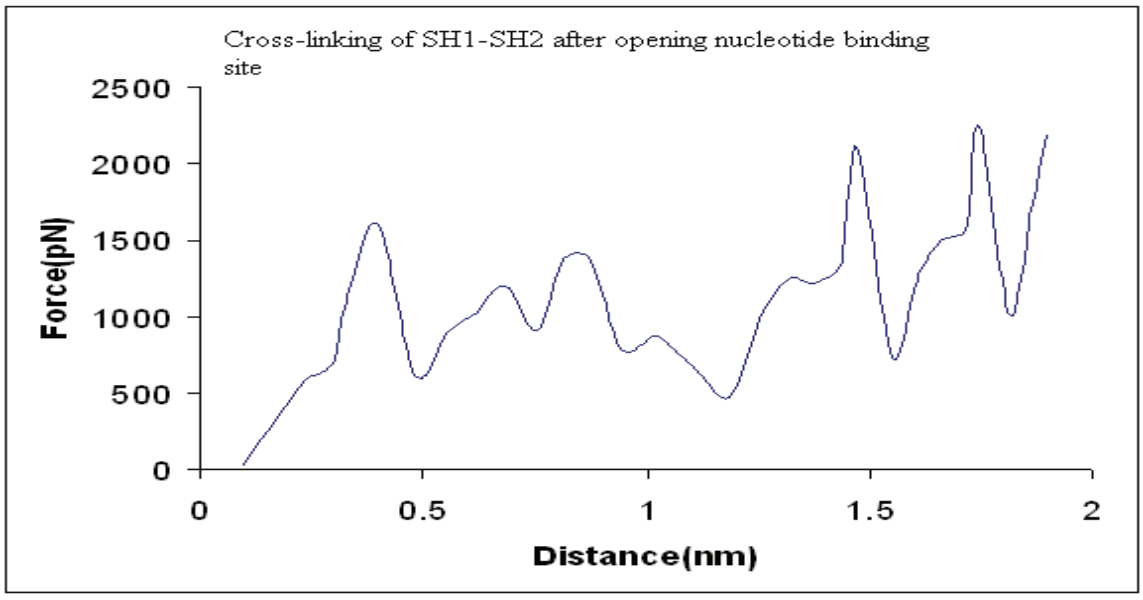

Fig. 7.10: Force-distance curve of crosslinking SH1-SH2 after opening nucleotide binding site. Force constraints with a spring constant of 16.7 N/m were used to open the nucleotide site SER 180 and SER 324 by changing a distance of 0.1nm increments followed by energy minimization and then crosslinking SH1-SH2 by changing a distance of 0.1nm increments followed by energy minimization.

Procedure for Opening/Closing the 50 kDa Cleft after Crosslinking SH1-SH2

The structure obtained from 100 ps of molecular dynamics simulations with the two cysteines crosslinked and at 0.3 nm with constraints attached to the two cysteines during simulation is taken as the starting structure . To this structure distance constraints were applied around the 50 kDa cleft to open/close the 50 kDa cleft. The distance constraints are incremented/decremented by 0.1nm during each energy minimization cycle and also the two cysteines are constrained and are kept at a distance of 0.3 nm. This process is continued till the 50 kDa cleft is opened/closed. This entire process takes 12 energy minimization cycles. Molecular dynamics simulations are done for 100 ps to the final structure with the two cysteines at 0.3 nm with constraints attached to the two cysteines during simulations. The final energy minimized structure is compared to the energy minimized structure of native myosin S1 after 100ps of simulation. The structure obtained after 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 200 ps of molecular dynamics simulations.

Simulations of the 50 kDa Cleft and Simulations at the Converter Region

The closing of the 50 kDa cleft structure with crosslinked cysteines (Fig. 7.11 A) does not show many changes when compared to closing the 50 kDa cleft (Fig 7.2 A). The strut gets more rigid when SH1 SH2 are crosslinked and the RMSD of the strut decreases (Table 3). Whereas, SH3 domain and converter increase in RMSD when the 50 kDa cleft is closed with crosslinked cysteines. This can be attributed to the constraints at the reactive cysteines. There is not a significant increase in work required to close the 50 kDa cleft in presence of crosslinked cysteines than to close the 50 kDa cleft in absence of crosslinked cysteines (Table. 2).

The opening of the 50 kDa cleft structure with crosslinked cysteines (Fig. 7.12 A) does not show significant changes when compared to opening the 50 kDa cleft in absence of crosslinked cysteines (Fig 7.1 A). The RMSD of the strut decreases on crosslinking but is not significant. The work required to open the cleft after crosslinking the reactive cysteines is less than the work required to open the 50 kDa cleft. The force-distance spectra of opening/closing the cleft in presence of crosslinked cysteines (Fig. 7.13) indicates that the force required to close the cleft is more than it is required to open in the presence of crosslinked cysteines.

The simulations performed for opening/closing the cleft in presence of crosslinked cysteines (Fig 7.11 B, 7.12 B) indicates that the lower 50 kDa domain shows more flexibility than the upper 50 kDa domain.

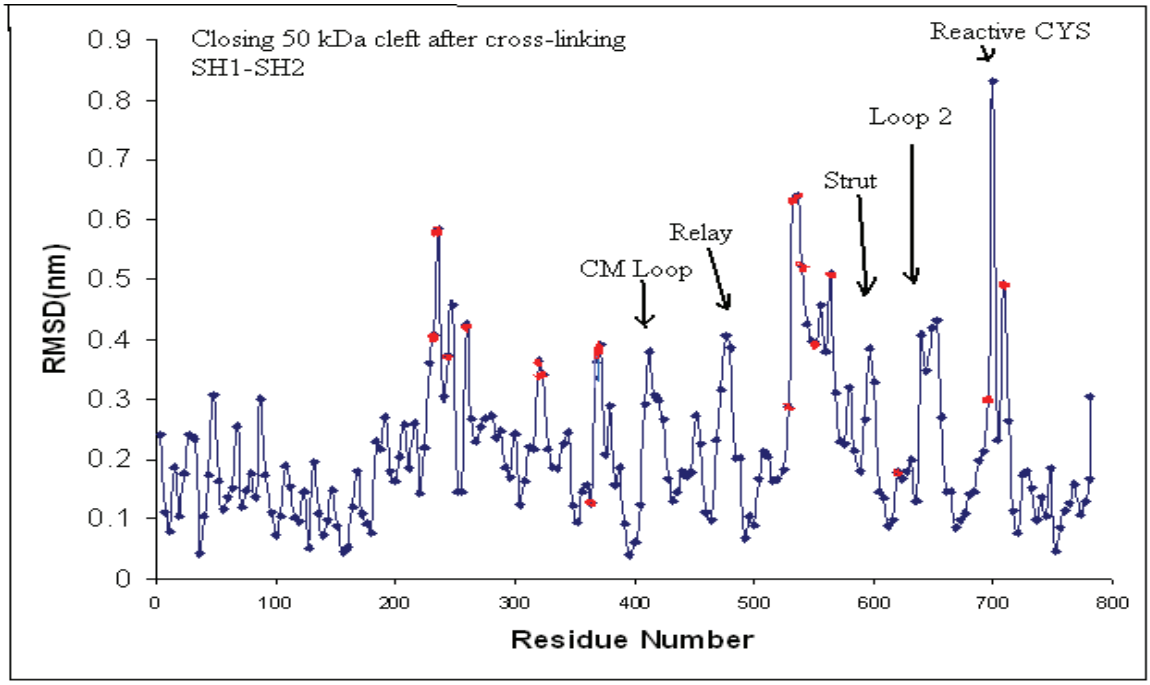

Fig. 7.11 A: Structural changes associated with closing of the 50kDa cleft after crosslinking the reactive cysteines SH1 and SH2. The endpoints of energy minimised structure is compared to the energy minimised structure of native myosin S1 after 100 ps of molecular dynamics simulations.

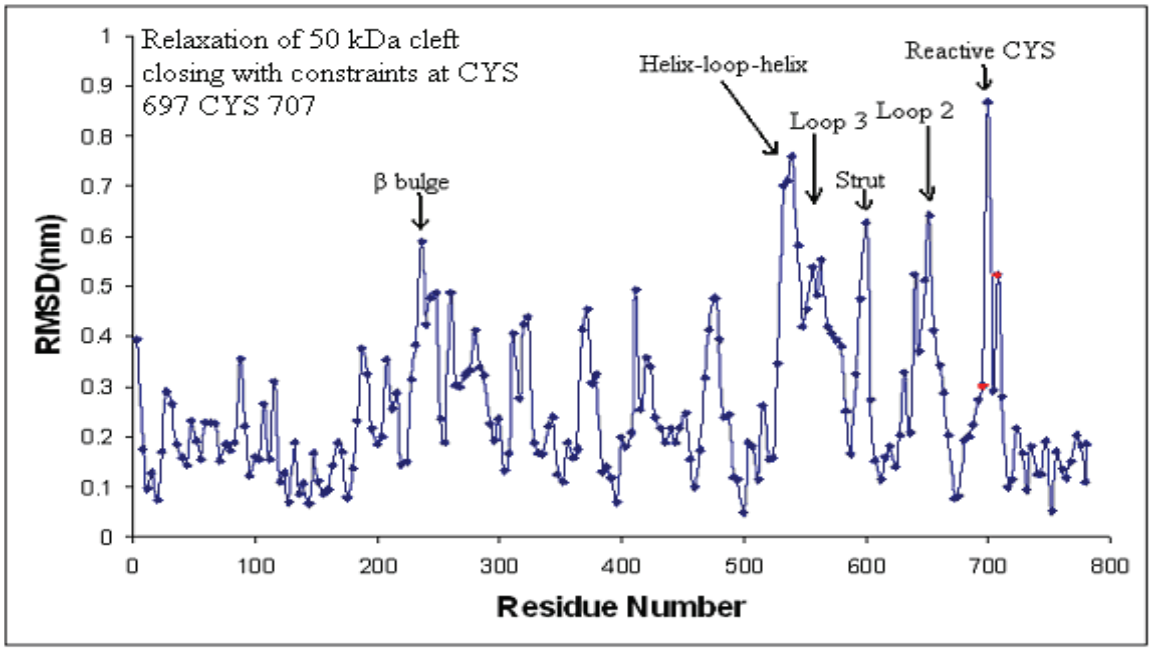

Fig 7.11 B: Structural changes associated with closing of the 50kDa cleft after crosslinking the reactive cysteines SH1 and SH2 . Constraints at SH1 and SH2 were present during molecular dynamics simulations. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 200 ps of molecular dynamics simulations.

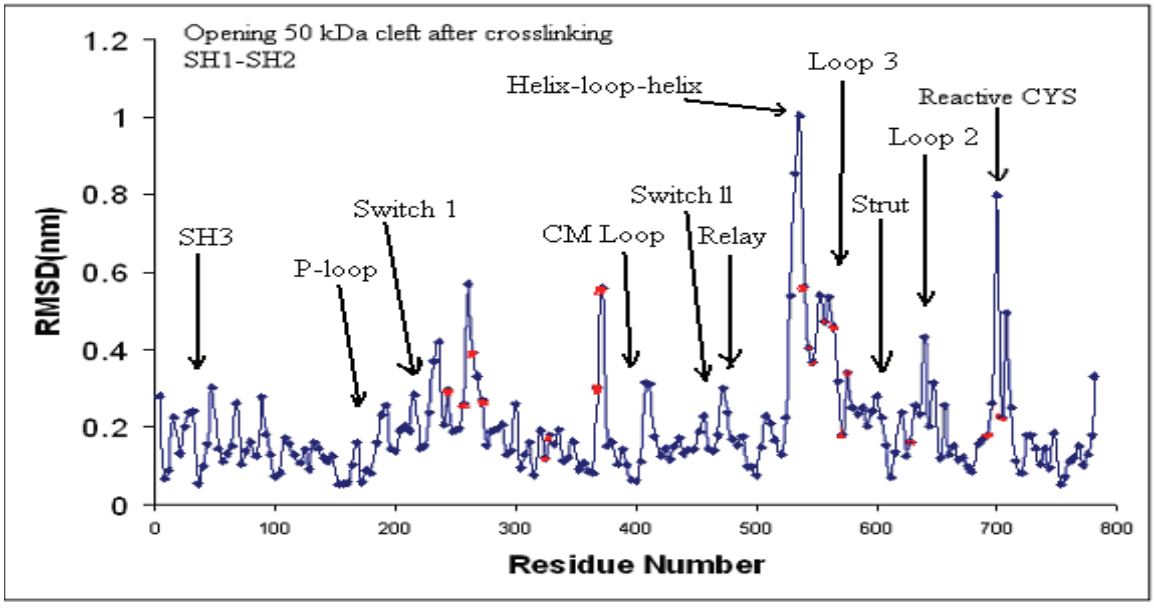

Fig 7.12 A: Structural changes associated with opening of the 50kDa cleft after crosslinking the reactive cysteines SH1 and SH2. The endpoints of energy minimised structure is compared to the energy minimised structure of native myosin S1 after 100 ps of molecular dynamics simulations.

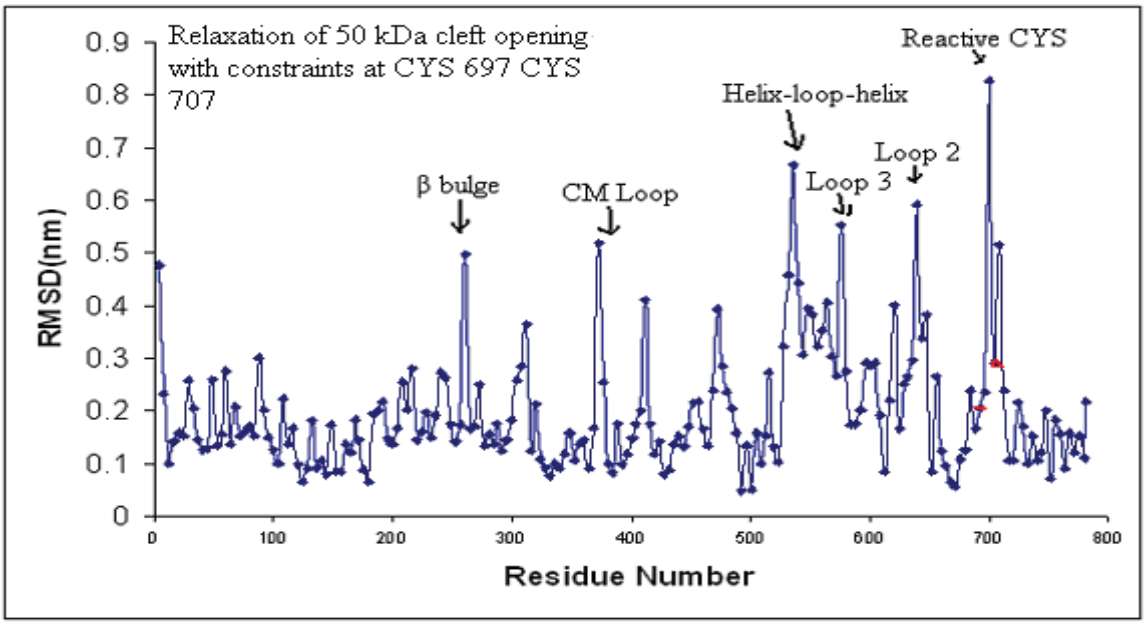

Fig 7.12 B: Structural changes associated with closing of the 50kDa cleft after crosslinking the reactive cysteines SH1 and SH2 . Constraints at SH1 and SH2 were present during molecular dynamics simulations. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 200 ps of molecular dynamics simulations.

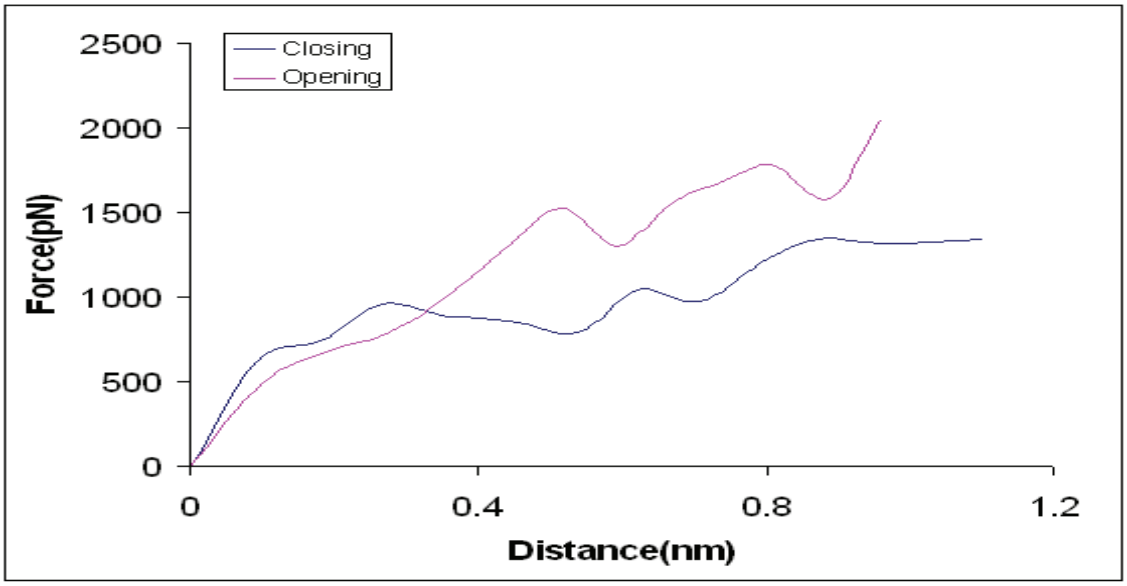

Fig. 7.13: Simulated closing and opening of the 50 kDa cleft after SH1and SH2 are crosslinked. Force constraints with a spring constant of 16.7 N/m were used to close and open the 50 kDa cleft by changing the target distance at 0.1 nm increments followed by energy minimization. A simulated force–distance curve was calculated based on the difference between the actual distance and the target distance after each energy minimization using the spring constant.

Procedure for Opening/Closing the 50 kDa Cleft with Nucleotide Binding Site Opened

The structure obtained from 100 ps of molecular dynamics simulations with constraints attached to the two serines and separated by 2.5 nm is taken as a starting structure for opening/closing the 50 kDa cleft with nucleotide binding site opened. To this structure distance constraints are applied around the 50 kDa cleft to close the 50 kDa cleft. The distance constraints are incremented/decremented by 0.1nm during each energy minimization cycle and also the two serines are constrained and are kept at a distance of 2.5 nm. This process is continued until the 50 kDa cleft is opened/closed. This entire process takes 12 energy minimization cycles. Molecular dynamics simulations are done for 100 ps to the final structure with the two serines at 0.3 nm with constraints attached to the two serines during simulations. The final energy minimized structure is compared to the energy minimized structure of native myosin S1 after 100 ps of simulation. The structure obtained from 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 200 ps of molecular dynamics.

Simulations of the 50 kDa Cleft with Nucleotide Binding Site Opened

The closing of the 50 kDa cleft structure with nucleotide binding site opened (Fig. 7.14 A) does not show many changes when compared to closing the 50 kDa cleft with nucleotide binding site closed (Fig 7.2 A). The significant changes are at the converter region and at the reactive cysteines which show an increase in RMSD when the cleft is closed with nucleotide binding site opened. The strut becomes more rigid when 50 kDa cleft is closed with nucleotide binding site opened. The work done to close the 50 kDa cleft with nucleotide binding site opened or closed is same.

The opening of the 50 kDa cleft structure with nucleotide binding site opened

(Fig. 7.15 A) does not show significant changes when compared to opening the 50 kDa cleft with nucleotide binding site closed (Fig 7.1 A). The significant changes are at the converter region and at the reactive cysteines which show an increase in RMSD when the 50 kDa cleft is opened with nucleotide binding site opened. The strut becomes more rigid when 50 kDa cleft is opened with nucleotide binding site open. The work required to open the 50 kDa cleft with nucleotide binding site open is more than the work required to open the 50 kDa cleft with nucleotide binding site closed. The work required to open the 50 kDa cleft with nucleotide binding site open is more than the work required to close the cleft with nucleotide binding site open (Fig. 7.16, Table. 2).

The simulations of the cleft for opening/closing with nucleotide binding site open indicate that the lower 50 kDa domain shows more flexibility than the upper 50 kDa domain (Fig 7.14 B, 7.15 B).

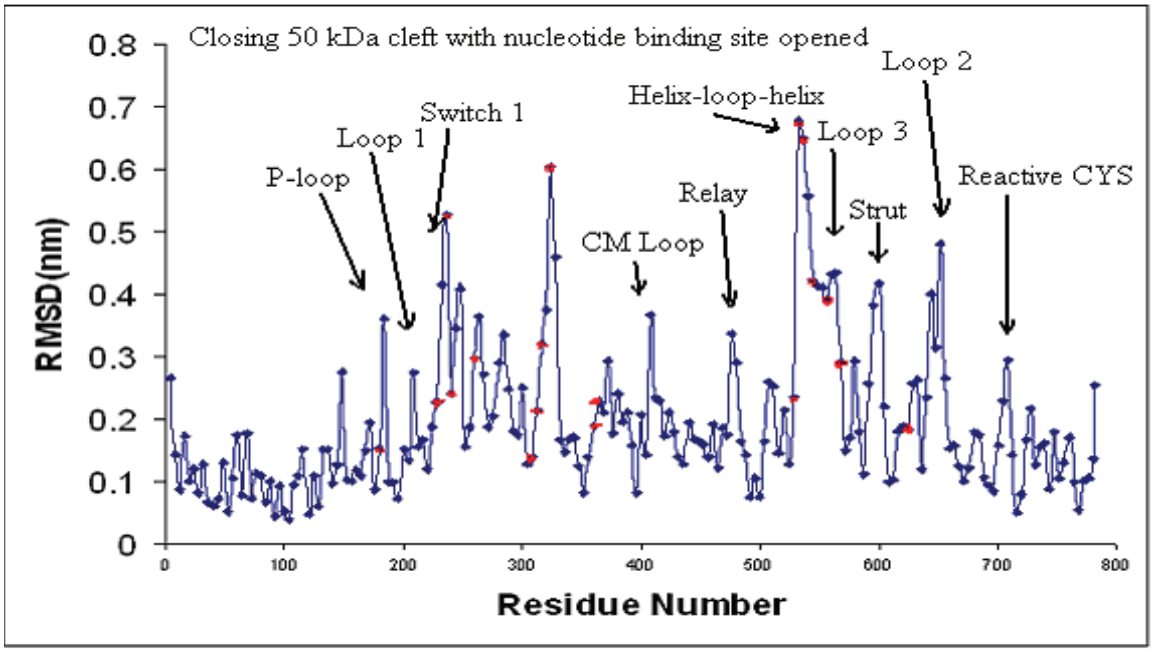

Fig. 7.14 A: Structural changes associated with closing of the 50 kDa cleft after opening nucleotide binding site. The endpoints of energy minimised structure is compared to the energy minimised structure of native myosin S1 after 100 ps of molecular dynamics simulations.

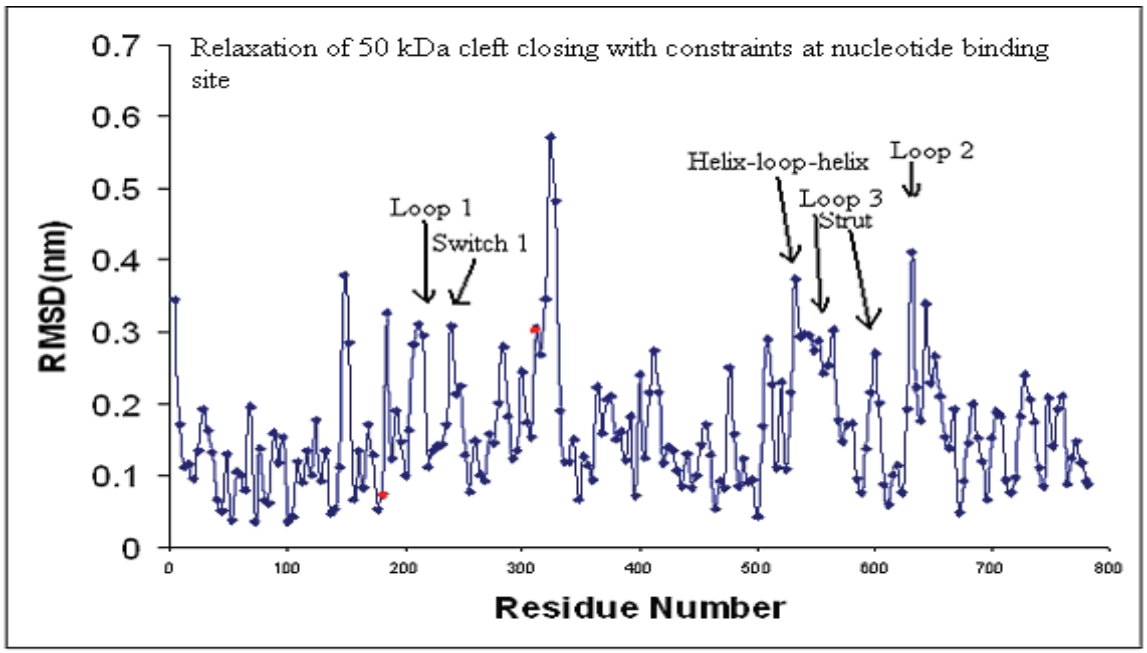

Fig. 7.14 B: Structural changes associated with closing of the 50kDa cleft after opening nucleotide site. Constraints at nucleotide site were present during molecular dynamics simulations. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 200 ps of molecular dynamics simulations.

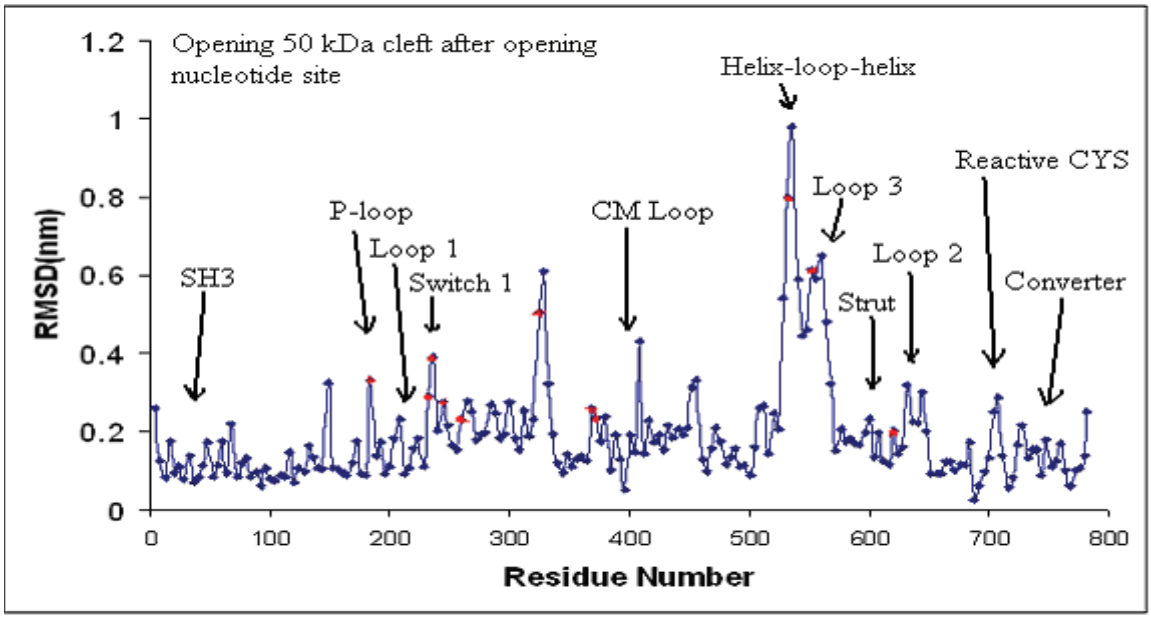

Fig. 7.15 A: Structural changes associated with opening of the 50 kDa cleft after opening the nucleotide binding site. The endpoints of energy minimised structure is compared to the energy minimised structure of native myosin S1 after 100 ps of molecular dynamics simulations.

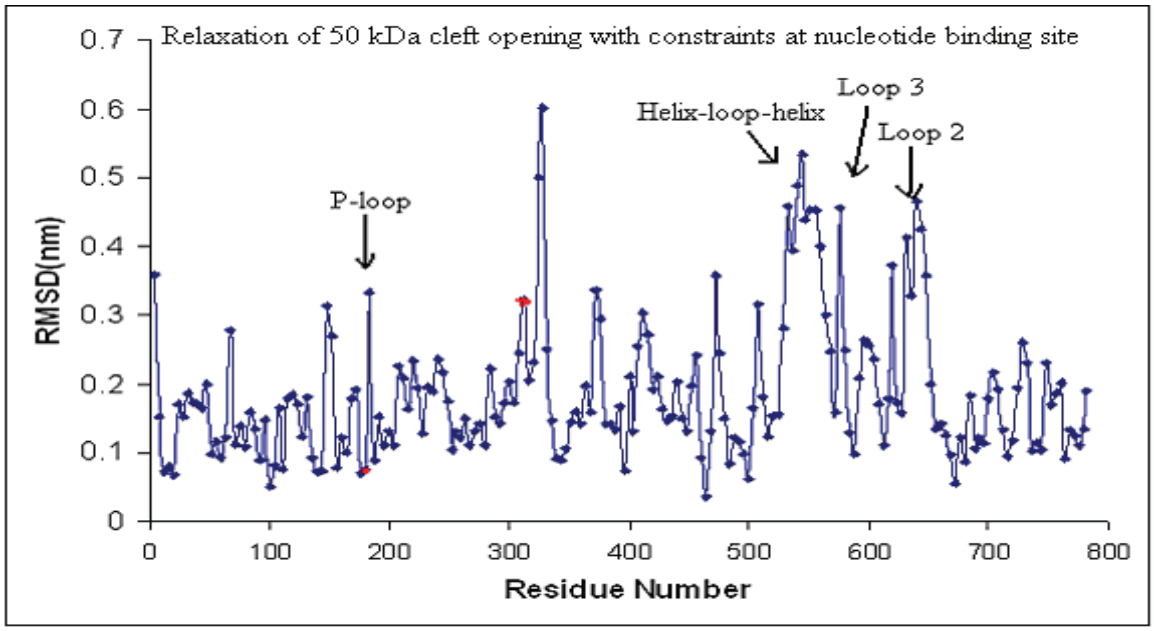

Fig. 7.15 B: Structural changes associated with opening of the 50kDa cleft after opening nucleotide site. Constraints at nucleotide site were present during molecular dynamics simulations. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 200 ps of molecular dynamics simulations.

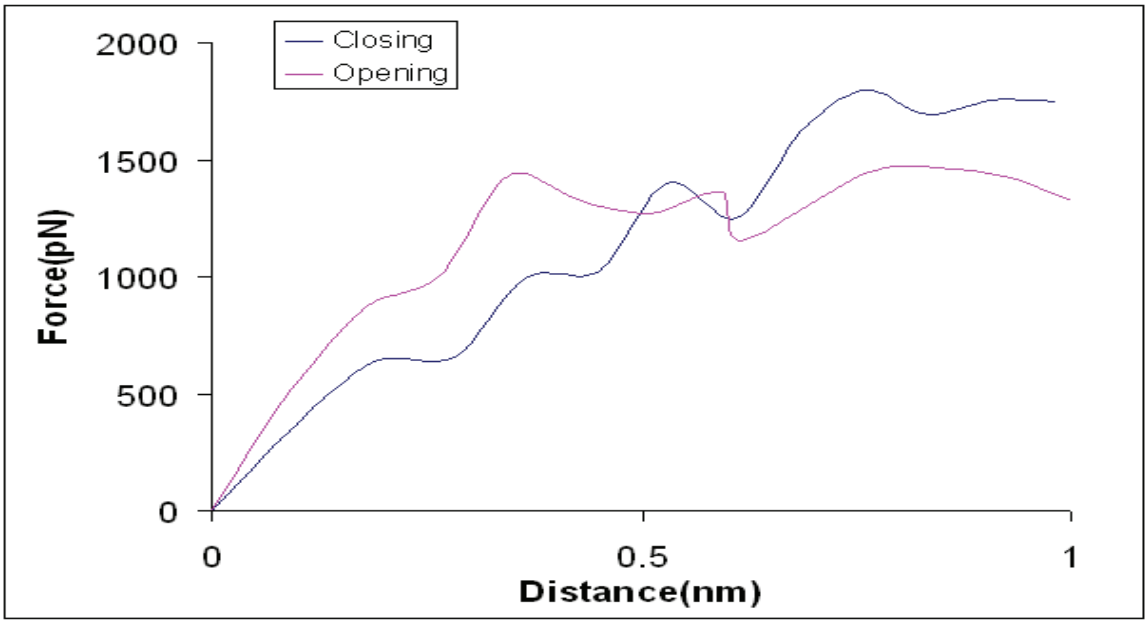

Fig. 7.16: Simulated closing and opening of the 50 kDa cleft after nucleotide binding site is opened. Force constraints with a spring constant of 16.7 N/m were used to close and open the 50 kDa cleft by changing the target distance at 0.1 nm increments followed by energy minimization. A simulated force–distance curve was calculated based on the difference between the actual distance and the target distance after each energy minimization using the spring constant.

Procedure for Opening/Closing the 50 kDa Cleft with Crosslinked SH1 SH2 and Nucleotide Binding site Opened

The structure obtained from molecular dynamics simulations for 100 ps with the two cysteines at 0.3 nm and the two serines at 2.5 nm with constraints attached to the two cysteines and the two serines during simulation is taken as the starting structure to open/close the 50 kDa cleft. To this structure distance constraints were applied around the 50 kDa cleft to open/close the 50 kDa cleft. The distance constraints are incremented/ decremented by 0.1 nm during each energy minimization cycle and also during the energy minimization the two serines are constrained and are kept at a distance of 2.5 nm and the two cysteines are also constrained and kept at a distance of 0.3 nm.

This process is continued until the 50 kDa cleft is opened/closed. This entire process takes 12 energy minimization cycles. Molecular dynamics simulations are done for 100 ps to the final structure with constraints attached to the two serines and the two cysteines during simulations. The final energy minimized structure is compared to the energy minimized structure of native myosin S1 after 200ps of simulation. The structure obtained from 100 ps of molecular dynamics simulations is compared to the native myosin S1 after 300 ps of molecular dynamics simulations.

Simulations of the 50 kDa Cleft, Nucleotide Binding Site and Simulations at the Converter Region

The closing of the 50 kDa cleft structure with crosslinked cysteines and nucleotide binding site opened (Fig. 7.17 A) does not show many changes when compared to closing the 50 kDa cleft with nucleotide binding site closed (Fig 7.2 A). The significant changes are at the converter region and at the reactive cysteines which show

an increase in RMSD when the cleft is closed with nucleotide binding site opened. The strut becomes more rigid when 50 kDa cleft is closed with crosslinked cysteines and nucleotide binding site opened but the change is not significant. The work required to close the 50 kDa cleft with crosslinked cysteines and nucleotide binding site opened is less than the work required to close the 50 kDa cleft with crosslinked cysteines and nucleotide binding site closed (Table. 2).

The opening of the 50 kDa cleft structure with crosslinked cysteines and nucleotide binding site opened (Fig. 7.18 A) does not show significant changes when compared to opening the 50 kDa cleft with nucleotide binding site closed (Fig. 7.1 A). The significant changes are at the converter region and at the reactive cysteines which show an increase in RMSD when the cleft is closed with nucleotide binding site opened. The strut becomes more rigid when 50 kDa cleft is opened with crosslinked cysteines and nucleotide binding site opened. The work required to open the 50 kDa cleft with crosslinked cysteines with nucleotide binding site opened is more than the work required to open the 50 kDa cleft with nucleotide binding site closed (Table. 2). The work required to open the cleft with crosslinked cysteines and nucleotide binding site opened is more than the work required to close the cleft with crosslinked cysteines and nucleotide binding site opened (Fig. 7.19, Table. 2).

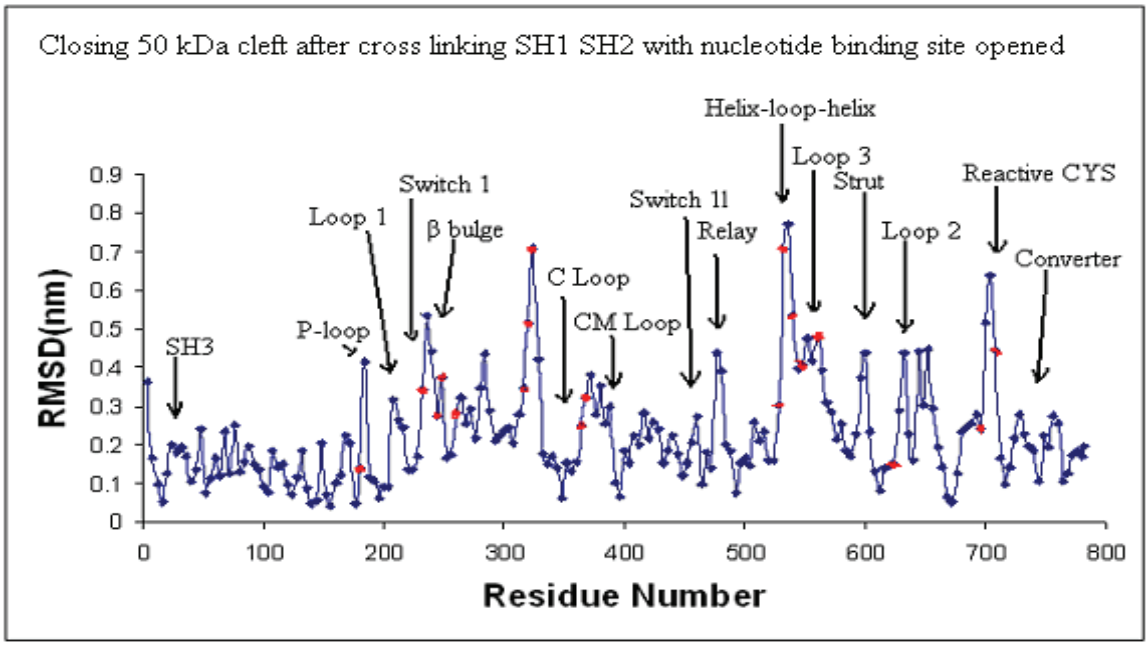

Fig. 7.17 A: Structural changes associated with closing the 50 kDa cleft with crosslinked cysteines and nucleotide binding site opened. The endpoints of energy minimized structure is compared to the energy minimized structure of native myosin S1 after 200 ps of molecular dynamics simulations.

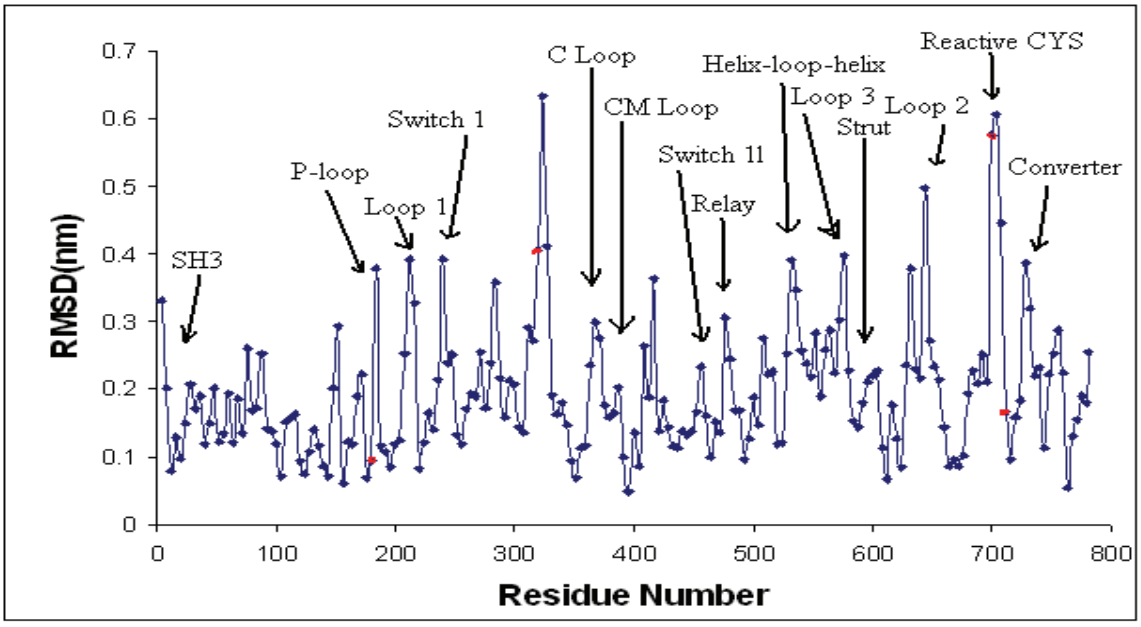

Fig. 7.17 B: Structural changes associated with closing of the 50 kDa cleft with crosslinked cysteines and nucleotide binding site opened. Constraints at reactive cysteines and nucleotide site were present during molecular dynamics simulations. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 300 ps of molecular dynamics simulations.

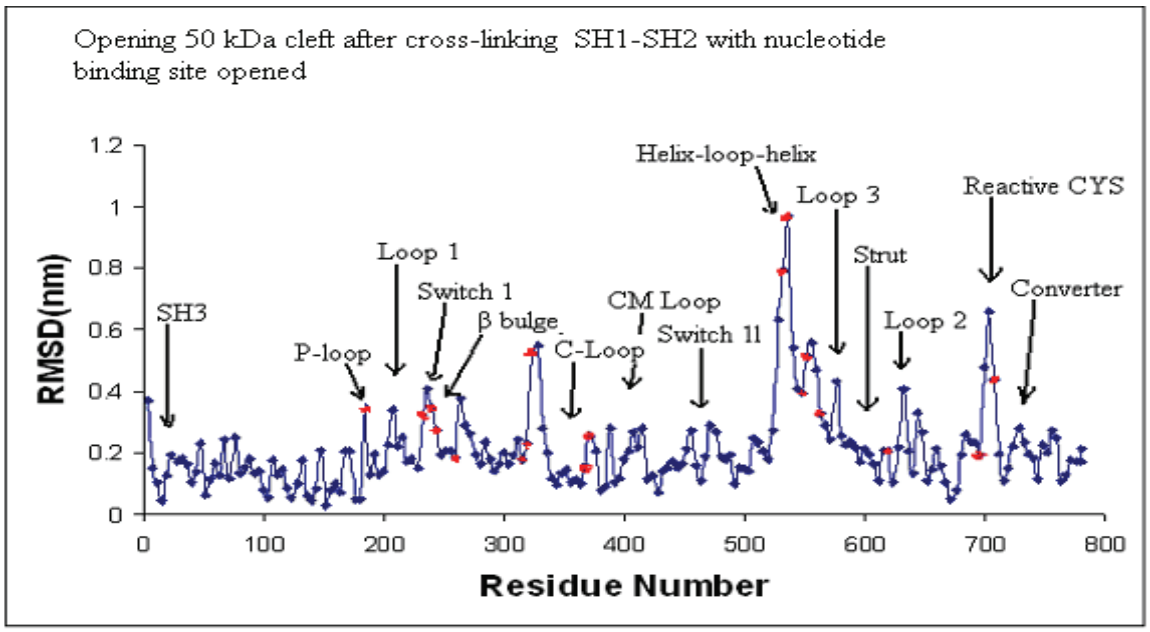

Fig. 7.18 A: Structural changes associated with opening the 50 kDa cleft with crosslinked cysteines and nucleotide binding site opened. The endpoints of energy minimised structure is compared to the energy minimised structure of native myosin S1 after 200 ps of molecular dynamics simulations.

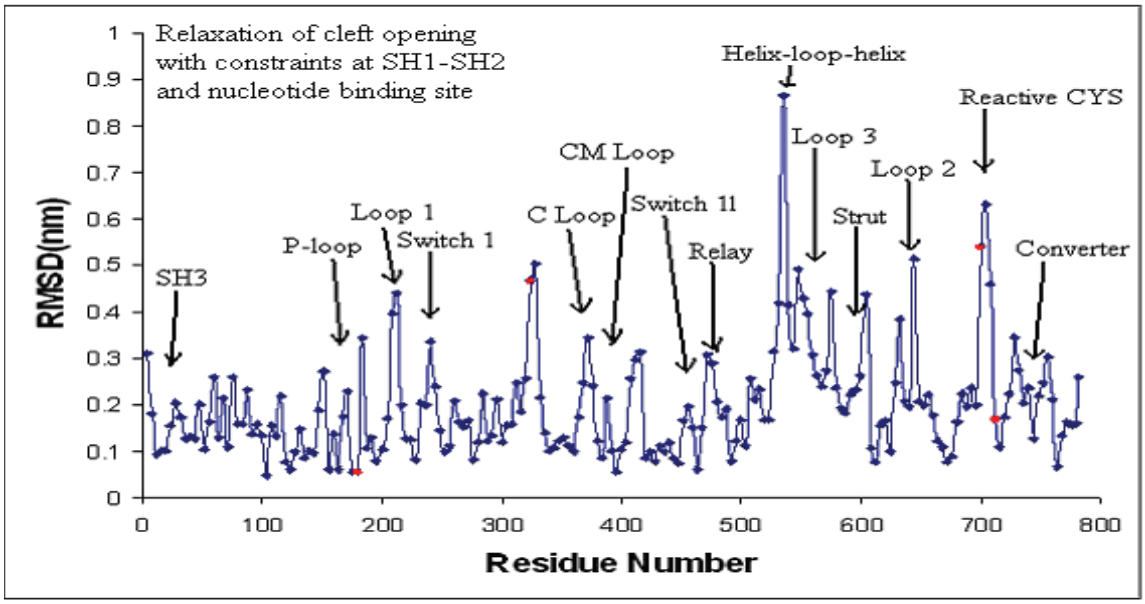

Fig. 7.18 B: Structural changes associated with opening of the 50 kDa cleft with crosslinked cysteines and nucleotide binding site opened. Constraints at reactive cysteines and nucleotide site were present during molecular dynamics simulations. The structure is compared to the molecular dynamic simulated structure of native myosin S1 after 300 ps of molecular dynamics simulations.

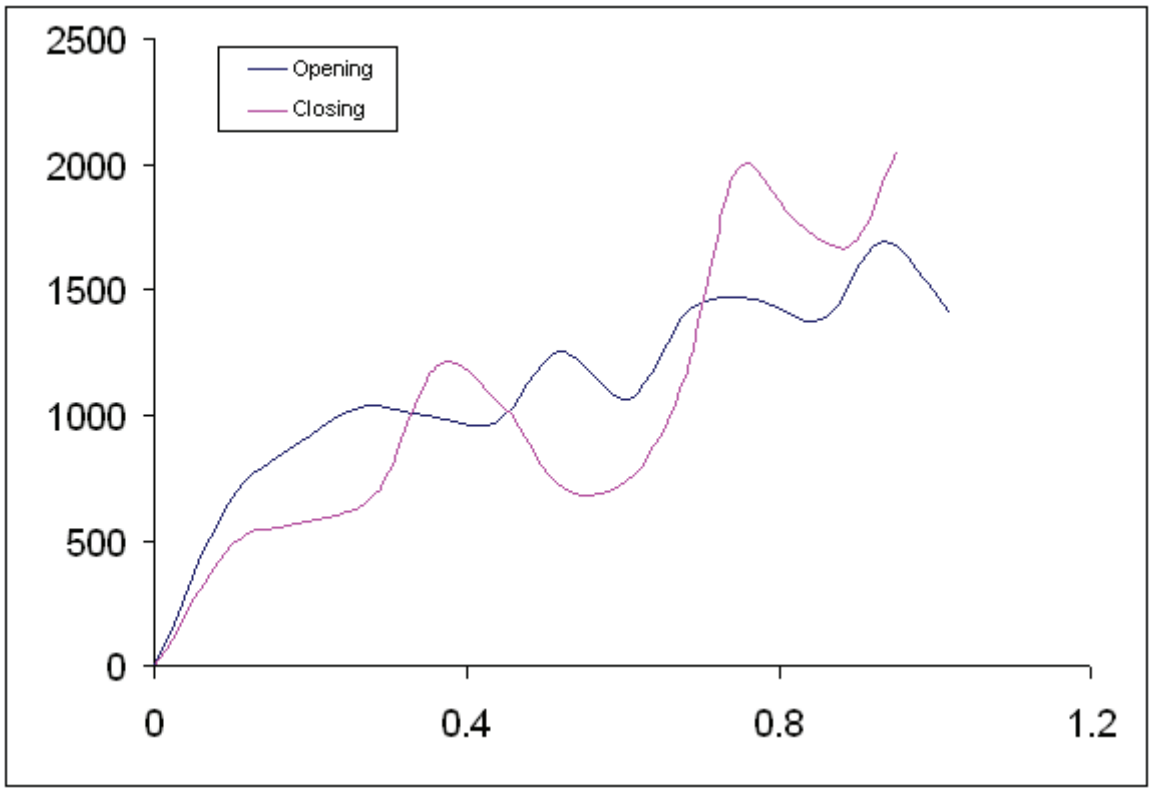

Fig. 7.19: Simulated closing and opening of the 50 kDa cleft with crosslinked cysteines with nucleotide binding site opened. Force constraints with a spring constant of 16.7 N/m were used to close and open the 50 kDa cleft by changing the target distance at 0.1 nm increments followed by energy minimization. A simulated force–distance curve was calculated based on the difference between the actual distance and the target distance after each energy minimization using the spring constant

Table 2: Work done for different conformations of myosin

CONFORMATIONAL STATES OF MYOSIN S1	WORK (zepto joules)
Opening 50 kDa cleft with nucleotide cleft closed	1038
Closing 50 kDa cleft with nucleotide cleft closed	1102
Opening 50 kDa cleft with nucleotide cleft opened	1147
Closing 50 kDa cleft with nucleotide cleft opened	1104
Opening 50 kDa cleft after cross linking SH1 SH2 with nucleotide cleft closed	978
Closing 50 kDa cleft after cross linking SH1 SH2 with nucleotide cleft closed	1133
Opening 50 kDa cleft after cross linking SH1 SH2 with nucleotide cleft opened	1137
Closing 50 kDa cleft after cross linking SH1 SH2 with nucleotide cleft opened	972
Cross linking of SH1 SH2 with nucleotide cleft closed	1733
Cross linking of SH1 SH2 with nucleotide cleft open	1917
Opening of nucleotide site	427

Table 3: RMSD of the strut during various conformational states of myosin S1

DIFFERENT CONFORMATIONAL STATES	RMSD OF THE STRUT (nano meters)
50 kDa cleft opening with nucleotide binding site closed	0.291
50 kDa cleft closing with nucleotide binding site closed	0.460
50 kDa cleft opening with nucleotide binding site opened	0.233
50 kDa cleft closing with nucleotide binding site opened	0.416
Cross linking of reactive cysteines with nucleotide binding site closed	0.0156
Opening of nucleotide binding site	0.0068
Cross linking reactive cysteines with nucleotide binding site opened	0.1312
Opening 50 kDa cleft after cross linking reactive cysteines	0.281
Closing 50 kDa cleft after cross linking reactive cysteines	0.3287
Opening 50 kDa cleft after cross linking reactive cysteines with nucleotide binding site opened	0.214
Closing 50 kDa cleft after cross linking reactive cysteines with nucleotide binding site opened	0.439

Based on the results from Table 2, it can be hypothesized that it is easier to open the 50 kDa cleft than close the 50 kDa cleft with nucleotide binding site closed. Whereas, with nucleotide binding site opened it becomes difficult to open the 50 kDa cleft but the closing of the 50 kDa cleft is easier. Cleft closing is independent of the opening or closing of the nucleotide binding site whereas cleft opening is dependent on the nucleotide binding site.

Opening the 50 kDa cleft is easier with crosslinked SH1 SH2 and nucleotide binding site closed .Whereas, closing the 50 kDa cleft becomes difficult with crosslinked SH1 SH2 and nucleotide binding site closed. Crosslinking the reactive cysteines aids in opening the 50 kDa cleft.

Opening the 50 kDa cleft is difficult with crosslinked SH1 SH2 and nucleotide binding site opened .Whereas, it becomes easier to close the 50 kDa cleft with crosslinked SH1 SH2 and nucleotide binding site opened. Crosslinking the reactive cysteines makes opening the 50 kDa cleft easier whereas opening of the nucleotide binding site makes it easier to close the 50 kDa cleft.

Based on the results from Table 3, it can be hypothesized that the strut is more mobile during the closing of 50 kDa cleft with nucleotide binding site closed.

Stochastic Dynamics Simulations of the S1 Motor Domain

To examine time dependent fluctuations in the motor domain during the picosecond time scale of fluorescence polarization measurements, stochastic dynamics of the skeletal muscle myosin motor domain were performed in a GBSA solvent model for up to 700 ps (Fig. 7.20). The atomic model of the chicken skeletal myosin [35] was modified by removing the light chain binding domain and adding predicted structures of the missing

surface loops using conformational searching techniques prior to the simulations. During the course of the dynamics simulations, there was substantial flexibility noted in some sequences of the motor domain (Fig. 7.20). The most flexible regions were the SH3 domain, the converter domain, the cardiomyopathy loop, and the surface loops 1, 2 & 3. The movement of lever arm is shown during the course of the dynamics simulations (Fig. 7.21). The strut sequence was among the more stable regions during the simulations yet it did demonstrate some degree of flexibility during the later time intervals of the simulation; however, the CY3 attachment site is in close proximity through space to loop 2 which might also influence the flexibility of CY3. Upon binding to actin, motions of actin, loop 2, or the 50 kDa cleft could potentially lead to the experimentally observed increase in CY3 flexibility from its initially less flexible attachment as judged by a fluorescence polarization of 0.42 for CY3-S1 compared to 0.30 for free CY3.

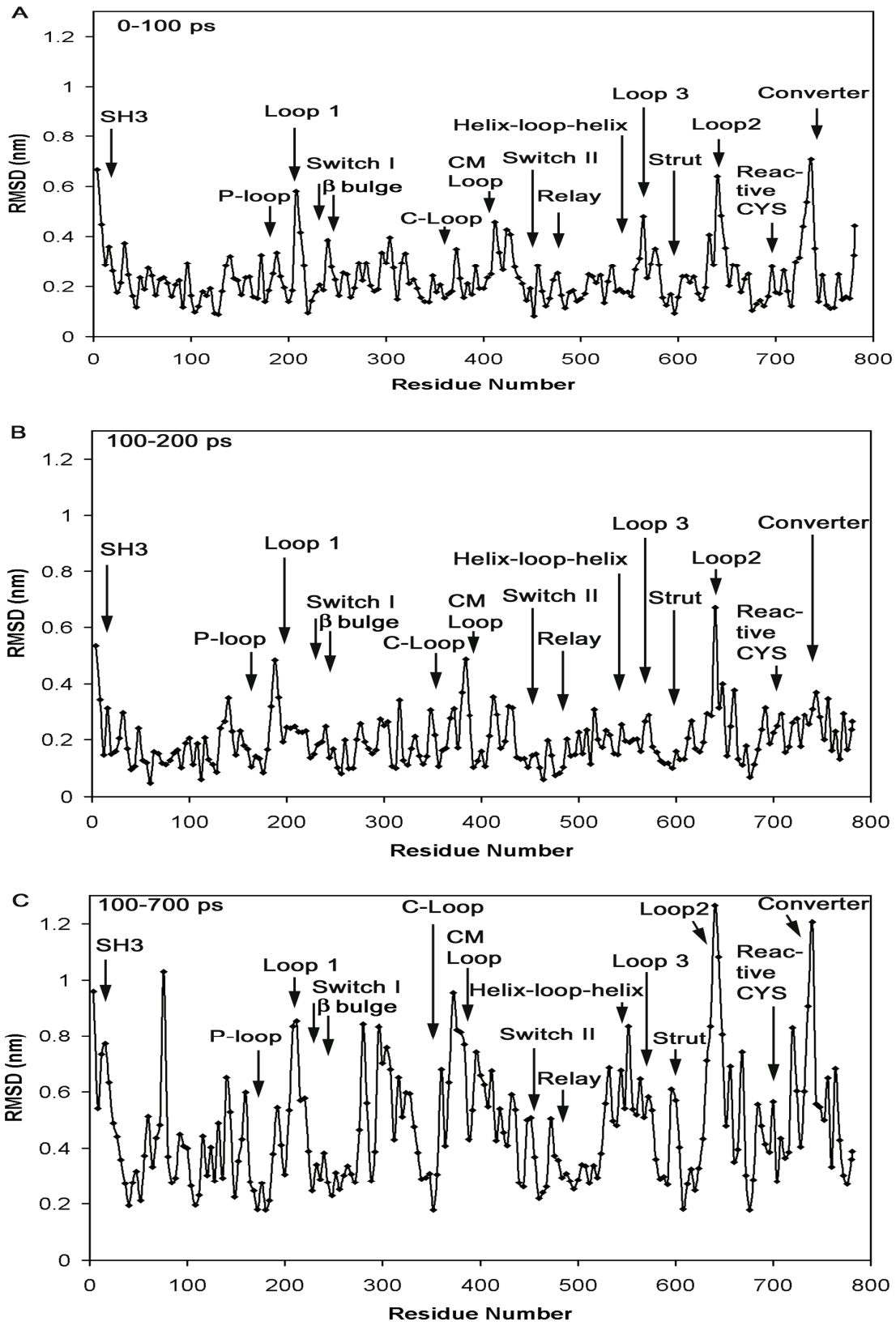

Fig. 7.20: Stochastic dynamics simulations of the skeletal muscle myosin motor domain.

To estimate the inherent flexibility of structural element in S1, dynamics simulations in a GBSA solvent model were performed for 700 ps on the energy minimized structure. The RMSD were compared for (**A**) the initial energy minimized structure with the energy minimized structure after 100 ps; (B) the energy minimized structure at 100 ps with the energy minimized structure after 200 ps; and (C) the energy minimized structure at 100 ps with the energy minimized structure after 700 ps. The motor domain from the skeletal muscle myosin atomic model, 1-792, was used for all computational simulations after the missing surface loops were reconstructed with LOOP algorithm and conformational searching routines in Macromodel. Key structural elements [113-115] are labeled: the N-terminal SH3-like domain (SH3), P-loop, loop 1, switch I, the bulge of the β-sheet (β bulge), cardiac loop (C-loop), cardiomyopathy loop (CM loop), switch II, relay helix (relay), helix-loop-helix, loop 3, strut, loop 2, reactive cysteines C697 & C707 (reactive CYS), and the converter region (converter).

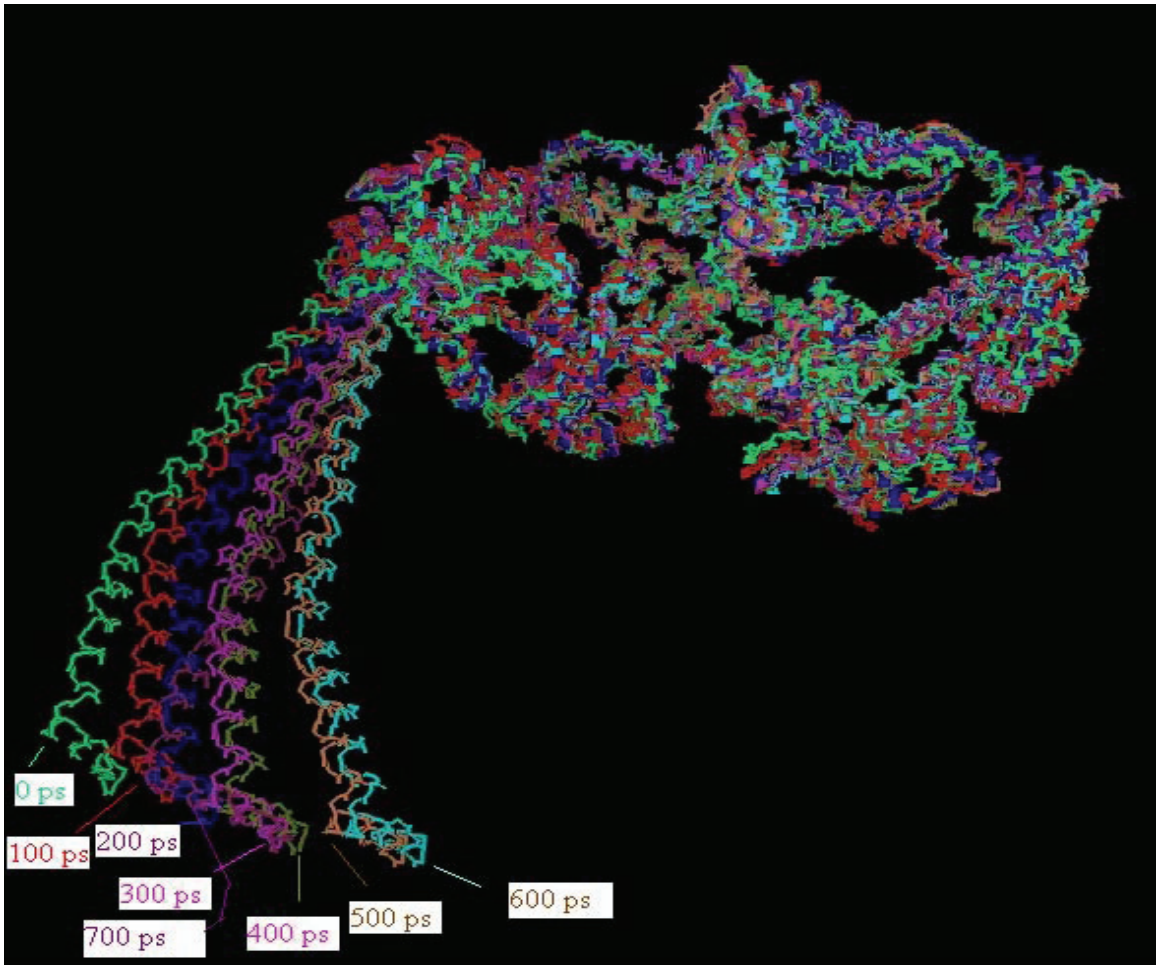

Fig. 7.21: Stochastic dynamics simulations of myosin S1 from 0 ps to 700 ps. The alpha-helix of the lever arm was reattached post-simulation to illustrate the impact of the various structural changes in the motor domain which leads to the movement of the lever arm.

CHAPTER 8

DISCUSSION

The interface between myosin and actin includes regions surrounding a cleft between myosin's upper and lower 50 kDa fragments. A sequence called the "strut" traverses this cleft and may well play an important role in their positioning during actin binding. Such a repositioning of elements around the cleft is thought to be intimately linked to more distant structural changes in the myosin's nucleotide cleft and neck regions. Structural analysis of the strut in myosin has not previously been performed by fluorescence spectroscopy. The development of a protocol to label and isolate skeletal myosin S1 with a single fluorescent probe at the strut sequence opens the door for spectroscopic studies of this key structure. To facilitate the interpretation of spectroscopic data, computational simulations of atomic models are becoming increasingly important, since fluorescent properties with lifetimes on the subnanosecond time scale reflect motions that can be simulated by dynamics calculations. Correlations of such motions with structural changes elsewhere in the molecule may also be implied by computational simulations. In this study, fluorescence measurements of CY3 attached to the strut of S1 are correlated with stochastic dynamics and molecular mechanics simulations of the cleft in myosin to reveal the impact of cleft motions on other key structures in the atomic model of the skeletal myosin motor domain.

The CY3 Label

The 50 kDa cleft of myosin contains a strut sequence that has been recently studied by mutagenesis in *Dictyostelium* [25-27, 111]. Changing the length of the strut sequence by insertion or deletion mutations strongly diminished its actin-activated

ATPase activity [25]. Similarly, mutating five of its residues to glycine without a change in the number of residues also reduced actin-activation of its ATPase activity [27]. However, the CY3 modification of the strut does not strongly affect the actin-activated S1 Vmax or the basal Mg2+ ATPase activity, still it does increase the Km much like the mutagenesis studies. It is likely that the CY3 modification does not greatly alter the length or stiffness of the strut. In addition, the labeled lysine is near the end of the strut.

With a five-carbon linker attaching the CY3 to a lysine, it is plausible that the dye is positioned away from the strut and instead interacts predominantly with surrounding residues. Support for this location comes from collisional quenching analysis that places the hydrophobic CY3 in close proximity to positively charged residues and on an accessible surface of the myosin [114]. Inspection of the atomic model indicates that there are a number of acidic residues in this region, but one on the neighbouring helixturn-helix motif contains hydrophobic and basic side chains, 540-544: CMFPK (rabbit). Indeed, K598 forms a salt bridge with E476 which makes it the less favoured labeling site relative to K600 that is in closest proximity to the helix-turn-helix and pointed toward the protein surface. Also, the lysines of the flexible loop 2 which is nearest K600 might contribute to the repulsion of the positively-charged Tl+ relative to I- during collisional quenching. Therefore, K600 is the most probable CY3 labelling site of the two strut lysines.

The mapped location of the CY3 dye places it near the actin binding site suggesting that actin binding may perturb the dye. The fluorescence polarization measurements confirmed that actin binding increased the mobility of the CY3, not unlike actin binding affects the fluorescence properties of W425 in smooth muscle myosin

[112]. Although it might have been speculated that the close proximity of actin would be more likely to reduce the mobility of the dye, it is possible that structural changes near the actomyosin interface could free the CY3 from a hydrophobic pocket on the surface of S1 giving rise to the increased rotation.

Intramolecular Communication

While none of computational simulations detected exceptional changes in the strut, the surrounding structures of the helix-loop-helix and loop 2 demonstrated significant flexibility. Some flexibility in the helix-loop-helix conformation was detected during the dynamics simulations after hundreds of picoseconds of simulation. These simulations suggest that the helix-turn-helix is not rigid. Although it has been hypothesized that the upper 50 kDa subdomain should be most mobile, these simulations suggest that the lower 50 kDa subdomain moves more readily in response to force constraints which is consistent with the possibility that significant structural rearrangements are necessary prior to a large movement of the upper 50 kDa subdomain.

The flexibility of loop 2 is evident in all of the simulations. This large surface loop is disordered in the original crystal structure and may serve to tether myosin to actin during the initial phases of actin binding [113]. There is potential for transient interactions of loop 2 with CY3, but more probable is a general electric shielding from the 5 lysines on loop 2 that likely contribute to the positive electric potential measured in the vicinity of CY3 on S1.

Some of the key sites in myosin do appear to be mechanically linked based on these simulations. The strut is a very rigid loop and needs a lot of work before it undergoes any significant conformational change. The strut loop is one of the three loops

connecting the upper and lower 50 kDa subdomains of myosin. Because the other two loops connecting the upper and lower 50 kDa subdomains are very flexible, the strut loop that has a stretched conformation and bulky side chains is expected to be rigid and keep the relative disposition of these two subdomains. The strut did exhibit modest changes in response to cleft movements. Closing of the 50 kDa cleft causes the strut to undergo conformational changes more than opening the 50 kDa cleft. Sutoh and his coworkers had also pointed out that the strut controls cleft closure in *Dictyostelium* myosin II [111]. Comparison of the 50 kDa cleft in open and closed states shows that the conformation of the strut loop varies between them and the relative disposition of the upper and lower 50K subdomains around the strut loop differ between the two states because of rotations of the two domains around this rigid loop. Although further changes in the relative orientation of the two domains are expected, the strut loop may set a limit for such changes.

In addition, the opening of the 50 kDa cleft seemed to correlate with some directed motions of the lever arm. It can be suggested that the narrow cleft that splits the central 50 kDa segment of the heavy chain provides not only the communication route between the nucleotide-binding pocket and actin but also transmits the conformational change necessary for movement. The exact interactions that cause these linkages are complex and will be the focus of future studies. There is hardly any movement in the strut during the crosslinking of the reactive cysteines SH1 and SH2 or opening of the nucleotide site. Crosslinking the reactive cysteines SH1 and SH2 or opening of the nucleotide site causes other loops to become more flexible but the strut becomes more rigid.

In summary, the novel labeling of the strut sequence of S1 with a fluorescent dye has been demonstrated and initial characterizations correlated well with predictions and simulations based on the location of the dye from peptide mapping. The CY3 labeling did not strongly affect Vmax of S1 which may make it a useful label for future studies. Interpretations of the simulations and experimental data collected support the view that the 50 kDa cleft is not easily mobile in the atomic model of skeletal S1. While limited motions of the cleft may occur readily, larger movements in the atomic model of the cleft likely require significant structural rearrangements relative to the atomic model to prepare it.

The events at the nucleotide site are tightly coupled to distant structural motifs, which are consistent with a broad range of experimental observations. Molecular simulations reveal how these mechanochemical coupling are reflected in the structural, energetic and dynamical features of the motor domain. Significant distortion of the actinbinding interface and the central β sheet has been observed to be induced by the opening of the nucleotide binding site in the simulations strongly suggesting the coupling between the opening of the nucleotide site and motion in the actin-binding interface. This hypothesis is consistent with the structural analysis of Geeves and Holmes [22], who proposed that during the powerstroke the opening of switch 1 induces the twisting of the central β sheet as well as the closure of the actin-binding cleft.

Crosslinking the reactive cysteines with nucleotide binding site opened required more work than crosslinking the reactive cysteines with the nucleotide binding site closed. It is well known from previous experiments that the SH1-SH2 crosslinking rate is affected by several factors, including the crosslinking span of the reagent, the presence of

ligands (nucleotides and/or actin) and temperature. The SH1-SH2 helix is transformed from a highly flexible structure in the ATP state into a rigid one in the prepower stroke state. When the nucleotide binding site is closed, the crosslinking is accelerated and the helix undergoes some conformational changes in order for SH1 and SH2 to come close to each other. Therefore, the work required to crosslink SH1 SH2 is less with nucleotide binding site closed and more to crosslink SH1 SH2 with nucleotide binding site opened. These results explain about helix melting or increased flexibility in the presence of nucleotides.

Switch I serves as one of the key sensors of the nucleotide binding site when 50 kDa cleft is open with crosslinked cysteines. Whereas, switch I and switch II are the key sensors of the nucleotide binding site when 50 kDa cleft is closed with crossed linked cysteines. It can be hypothesized that switch II is the one which is causing the complete reversal of energy when SH1 SH2 is crosslinked with nucleotide binding site opened and closed.

The hinge residues are distributed over several key inter-subdomain joints including the nucleotide-binding pocket, the relay helix, the SH1 helix, the strut between the upper 50 kDa and the lower 50 kDa subdomains, which is consistent with their hypothesized roles in modulating functionally relevant inter-subdomain conformational changes. These simulations suggest plausible linkages between several of the key motor domain structures; however, the strut sequence is largely unaffected.

Conclusion

In conclusion, myosin is a complicated protein which goes through several conformational changes upon its binding to actin and the power stroke to form a muscle

contraction. Molecular dynamic simulations clearly indicate that the myosin motor domain has intrinsic flexibility. In summary, a comparison of the mobility of the strut indicates that it is essentially a rigid body. The strut remains rigid under conditions known to affect the global motion of the myosin head and behaves as a single body. The strut is affected mostly by closure of the 50 kDa cleft in the presence of the nucleotide. It is easier to open the 50 kDa cleft than close the 50 kDa cleft in presence of a nucleotide. Whereas, in the absence of the nucleotide it becomes difficult to open the cleft but has no effect on cleft closure. Cleft closure is independent on the nucleotide whereas opening the 50 kDa cleft is dependent on the nucleotide. Crosslinking of reactive cysteines in the absence of a nucleotide requires more work than in the presence of a nucleotide. Crosslinking the reactive cysteines makes opening the 50 kDa cleft easier in presence of a nucleotide and closing difficult. Whereas, crosslinking the reactive cysteines in the absence of the nucleotide makes it easier to close the 50 kDa cleft and difficult to open.

The lever arm is fairly rigid and requires considerable amount of work to make it move. For a motor, which generates strain in one domain and transmits it to other domain there should be regions of flexibility between the domains while the domain transmitting the force, the lever arm should be fairly rigid. The lever arm moves during opening of 50 kDa cleft and there is hardly any change when the 50 kDa cleft is closed.

The labeling methodology developed herein has potentially general applicability. By using various combinations of molar ratios of protein and different fluorescent labels, it is possible to isolate singly labeled protein of any region of interest and study the conformational and structural changes of the labeled region of a protein and hence gain

more understanding about the mechanism of mutated proteins.

Conformational changes in proteins are collective phenomena. To fully understand the mechanism of mechanochemical coupling in myosin, it is important to reveal residues and interactions that dictate the coordination among ATP hydrolysis, opening and closing of active site, converter rotation and changes in the actin-binding site, opening and closing of 50 kDa cleft. This can be obtained by complex simulations and experimental methods.

The results obtained in the present study are similar to the results obtained from previous experimental studies from other groups on crosslinking reactive cysteines in presence and absence of nucleotide, the role of strut in cleft closure, rigidity of the strut, the force(work) required for a specific conformational change. Some of the experiments like opening and closing of 50 kDa cleft, order and magnitude of structural changes, the flexibility of critical regions, lever arm movement when cleft is opened or closed cannot be solved using experimental methods. Computational simulations helped in understanding these complex conformational changes. Myosin is a prototypical molecular motor therefore, insights and computational protocols established for myosin might be transferable to other molecular motor systems. In addition there is a rich experimental background for myosin, including diverse structural mutagenesis, and spectroscopic and motility characterizations, which makes it possible to extensively validate the computational analyses. Myosin II is uniquely suited for theoretical and computational analysis due to the availability of multiple x-ray structures at high resolution and the rich experimental background. A full determination of how this protein acts could better help us understand how defects in any of its regions could result in disease or death. In future,

the results reported here with additional simulations of the lever arm with both the regulatory light chain and essential light chain will facilitate in development of an effective model that can simulate the entire functional cycle of myosin.

.

REFERENCES

1) Vale, R. D., and Milligan, R. A. (2000) The way things move: looking under the hood of molecular motor proteins, *Science (New York, N.Y 288*, 88-95.

2) Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K., Jr. (1997) Direct observation of the rotation of F1-ATPase, *Nature 386*, 299-302.

3) Wang, M. D., Schnitzer, M. J., Yin, H., Landick, R., Gelles, J., and Block, S. M. (1998) Force and velocity measured for single molecules of RNA polymerase, *Science (New York, N.Y 282*, 902-907.

4) Doering, C., Ermentrout, B., and Oster, G. (1995) Rotary DNA motors, *Biophysical journal 69*, 2256-2267.

5) Wilson, K. S., and Noller, H. F. (1998) Mapping the position of translational elongation factor EF-G in the ribosome by directed hydroxyl radical probing, *Cell 92*, 131-139.

6) Mixon, M. B., Lee, E., Coleman, D. E., Berghuis, A. M., Gilman, A. G., and Sprang, S. R. (1995) Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis, *Science (New York, N.Y 270*, 954-960.

7) Vale, R. D. (1996) Switches, latches, and amplifiers: common themes of G proteins and molecular motors, *The Journal of cell biology 135*, 291-302.

8) Berg, J. S., Powell, B. C., and Cheney, R. E. (2001) A millennial myosin census, *Molecular biology of the cell 12*, 780-794.

9) Thompson, R. F., and Langford, G. M. (2002) Myosin superfamily evolutionary history, *The Anatomical record 268*, 276-289.

10) Richards , T. A, and Cavalier-Smith, T.(2005)Myosin domain evolution and the primary divergence of eukaryotes, *Nature 436*,1113-1118.

11) Vale,R. D (2003)The molecular motor toolbox for intercellular transport, *Cell 112,* 467-480.

12) Pollard, T. D., and Korn, E. D. (1973) Acanthamoeba myosin. I. Isolation from Acanthamoeba castellanii of an enzyme similar to muscle myosin, *The Journal of biological chemistry 248*, 4682-4690.

13) Higuchi, H., and Endow, S. A. (2002) Directionality and processivity of molecular motors, *Current opinion in cell biology 14*, 50-57.

14) Mehta, A. D., Rock, R. S., Rief, M., Spudich, J. A., Mooseker, M.

S., and Cheney, R.E. (1999) Myosin-V is a processive actin-based motor, *Nature 400*, 590-593.

15) Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. (2003) Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization, *Science (New York, N.Y 300*, 2061-2065.

16) Friedman, T. B., Sellers, J. R., and Avraham, K. B. (1999) Unconventional myosins and the genetics of hearing loss, *American journal of medical genetics 89*, 147-157.

17) Engle, L. J., and Kennett, R. H. (1994) Cloning, analysis, and chromosomal localization of myoxin (MYH12), the human homologue to the mouse dilute gene, *Genomics 19*, 407-416.

18) Watkins, H., Rosenzweig, A., Hwang, D. S., Levi, T., McKenna, W., Seidman, C. E., and Seidman, J. G. (1992) Characteristics and prognostic implications of myosin missense mutations in familial hypertrophic cardiomyopathy, *The New England journal of medicine 326*, 1108-1114.

19) Toyoshima, Y. Y., Kron, S. J., McNally, E. M., Niebling, K. R., Toyoshima, C., and Spudich, J. A. (1987) Myosin subfragment-1 is sufficient to move actin filaments in vitro, *Nature 328*, 536-539.

20) Rayment, I., Rypniewski, W. R., Schmidt-Base, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G., and Holden, H. M. (1993) Three-dimensional structure of myosin subfragment-1: a molecular motor, *Science (New York, N.Y 261*, 50-58.

21) Xie, X., Harrison, D. H., Schlichting, I., Sweet, R. M., Kalabokis, V. N., Szent-Gyorgyi, A. G., and Cohen, C. (1994) Structure of the regulatory domain of scallop myosin at 2.8 A resolution, *Nature 368*, 306-312.

22) Geeves, M. A., and Holmes, K. C. (1999) Structural mechanism of muscle contraction, *Annual review of biochemistry 68*, 687-728.

23) Cooke, R. (1986) The mechanism of muscle contraction, *CRC critical reviews in biochemistry 21*, 53-118.

24) Holmes, K. C. (1997) The swinging lever-arm hypothesis of muscle contraction, *Curr Biol 7*, R112-118.

25) Sasaki, N., Ohkura, R., and Sutoh, K. (2000) Insertion or deletion of a single residue in the strut sequence of Dictyostelium myosin II abolishes strong binding to actin, *The Journal of biological chemistry 275*, 38705-38709.

26) Zheng, W., and Brooks, B. (2005) Identification of dynamical correlations within

the myosin motor domain by the normal mode analysis of an elastic network model, *Journal of molecular biology 346*, 745-759.

27) Fujita-Becker, S., Reubold, T. F., and Holmes, K. C. (2006) The actin-binding cleft: functional characterisation of myosin II with a strut mutation, *Journal of muscle research and cell motility 27*, 115-123.

28) Spudich, J. A. (1994) How molecular motors work, *Nature 372*, 515-518

29) Uyeda, T. Q., Abramson, P. D., and Spudich, J. A. (1996) The neck region of the myosin motor domain acts as a lever arm to generate movement, *Proceedings of the National Academy of Sciences of the United States of America 93*, 4459-4464.

30) Bagshaw, C. R., and Trentham, D. R. (1974) The characterization of myosin-product complexes and of product-release steps during the magnesium ion-dependent adenosine triphosphatase reaction, *The Biochemical journal 141*, 331-349.

31) Werber, M. M., Szent-Gyorgyi, A. G., and Fasman, G. D. (1972) Fluorescence studies on heavy meromyosin-substrate interaction, *Biochemistry 11*, 2872-2883.

32) Pollard, T. D., and Ostap, E. M. (1996) The chemical mechanism of myosin-I: implications for actin-based motility and the evolution of the myosin family of motor proteins, *Cell structure and function 21*, 351-356.

33) Hanson, J. (1967). Axial period of actin filaments. *Nature, 21, 353-356*

34) Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., and Holmes, K. C. (1990) Atomic structure of the actin:DNase I complex, *Nature 347*, 37-44.

35) Rayment, I., Holden, H. M., Whittaker, M., Yohn, C. B., Lorenz, M., Holmes, K. C., and Milligan, R. A. (1993) Structure of the actin-myosin complex and its implications for muscle contraction, *Science (New York, N.Y 261*, 58-65.

36) Milligan, R. A. (1996) Protein-protein interactions in the rigor actomyosin complex, *Proceedings of the National Academy of Sciences of the United States of America 93*, 21-26.

37) Sata, M., and Ikebe, M. (1996) Functional analysis of the mutations in the human cardiac beta-myosin that are responsible for familial hypertrophic cardiomyopathy. Implication for the clinical outcome, *The Journal of clinical investigation 98*, 28662873.

38) Fujita, H., Sugiura, S., Momomura, S., Omata, M., Sugi, H., and Sutoh, K. (1997) Characterization of mutant myosins of Dictyostelium discoideum equivalent to human familial hypertrophic cardiomyopathy mutants. Molecular force level of

mutant myosins may have a prognostic implication, *The Journal of clinical investigation 99*, 1010-1015.

39) Rayment, I., Holden, H. M., Sellers, J. R., Fananapazir, L., and Epstein, N. D. (1995) Structural interpretation of the mutations in the beta-cardiac myosin that have been implicated in familial hypertrophic cardiomyopathy, *Proceedings of the National Academy of Sciences of the United States of America 92*, 3864-3868.

40) Dominguez, R., Freyzon, Y., Trybus, K. M., and Cohen, C. (1998) Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state, *Cell 94*, 559-571.

41) Houdusse, A., Kalabokis, V. N., Himmel, D., Szent-Gyorgyi, A. G., and Cohen, C. (1999) Atomic structure of scallop myosin subfragment S1 complexed with MgADP: a novel conformation of the myosin head, *Cell 97*, 459-470.

42) Taylor, E. W. (1979) Mechanism of actomyosin ATPase and the problem of muscle contraction, *CRC critical reviews in biochemistry 6*, 103-164.

43) Mornet, D., Pantel, P., Audemard, E., and Kassab, R. (1979) The limited tryptic cleavage of chymotryptic S-1: an approach to the characterization of the actin site in myosin heads, *Biochemical and biophysical research communications 89*, 925-932.

44) Duke, J., Takashi, R., Ue, K., and Morales, M. F. (1976) Reciprocal reactivities of specific thiols when actin binds to myosin, *Proceedings of the National Academy of Sciences of the United States of America 73*, 302-306.

45) Conibear, P. B., Bagshaw, C. R., Fajer, P. G., Kovacs, M., and Malnasi-Csizmadia, A. (2003) Myosin cleft movement and its coupling to actomyosin dissociation, *Nature structural biology 10*, 831-835.

46) Jarcho, J. A., McKenna, W., Pare, J. A., Solomon, S. D., Holcombe, R. F., Dickie, S., Levi, T., Donis-Keller, H., Seidman, J. G., and Seidman, C. E. (1989) Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1, *The New England journal of medicine 321*, 1372-1378.

47) Maron, B. J., Roberts, W. C., Edwards, J. E., McAllister, H. A., Jr., Foley, D. D., and Epstein, S. E. (1978) Sudden death in patients with hypertrophic cardiomyopathy: characterization of 26 patients with functional limitation, *The American journal of cardiology 41*, 803-810.

48) Maron, B. J., Epstein, S. E., and Roberts, W. C. (1986) Causes of sudden death in competitive athletes, *Journal of the American College of Cardiology 7*, 204-214.

49) Watkins, H., Thierfelder, L., Anan, R., Jarcho, J., Matsumori, A., McKenna, W., Seidman, J. G., and Seidman, C. E. (1993) Independent origin of identical beta

cardiac myosin heavy-chain mutations in hypertrophic cardiomyopathy, *American journal of human genetics 53*, 1180-1185.

50) Dausse, E., and Schwartz, K. (1993) Genetic heterogeneity of familial hypertrophic cardiomyopathy, *Neuromuscul Disord 3*, 483-486.

51) Marian, A. J., Mares, A., Jr., Kelly, D. P., Yu, Q. T., Abchee, A. B., Hill, R., and Roberts, R. (1995) Sudden cardiac death in hypertrophic cardiomyopathy. Variability in phenotypic expression of beta-myosin heavy chain mutations, *European heart journal 16*, 368-376.

52) Epstein, N. D., Cohn, G. M., Cyran, F., and Fananapazir, L. (1992) Differences in clinical expression of hypertrophic cardiomyopathy associated with two distinct mutations in the beta-myosin heavy chain gene. A 908Leu----Val mutation and a 403Arg----Gln mutation, *Circulation 86*, 345-352.

53) Morita, F. (1967) Interaction of heavy meromyosin with substrate. I. Difference in ultraviolet absorption spectrum between heavy meromyosin and its Michaelis-Menten complex, *The Journal of biological chemistry 242*, 4501-4506.

54) Trentham, D. R., Eccleston, J. F., and Bagshaw, C. R. (1976) Kinetic analysis of ATPase mechanisms, *Quarterly reviews of biophysics 9*, 217-281.

55) Sutoh, K., and Lu, R. C. (1987) Identification of two segments, separated by approximately 45 kilodaltons, of the myosin subfragment 1 heavy chain that can be cross-linked to the SH-1 thiol, *Biochemistry 26*, 4511-4516.

56) Huston, E. E., Grammer, J. C., and Yount, R. G. (1988) Flexibility of the myosin heavy chain: direct evidence that the region containing SH1 and SH2 can move 10 A under the influence of nucleotide binding, *Biochemistry 27*, 8945-8952.

57) Seidel, J. C., Chopek, M., and Gergely, J. (1970) Effect of nucleotides and pyrophosphate on spin labels bound to S1 thiol groups of myosin, *Biochemistry 9*, 3265-3272.

58) Wells, C., and Bagshaw, C. R. (1984) The characterization of vanadate-trapped nucleotide complexes with spin-labelled myosins, *Journal of muscle research and cell motility 5*, 97-112.

59) Barnett, V. A., and Thomas, D. D. (1987) Resolution of conformational states of spin-labeled myosin during steady-state ATP hydrolysis, *Biochemistry 26*, 314-323.

60) Walker, M., and Trinick, J. (1988) Visualization of domains in native and nucleotide-trapped myosin heads by negative staining, *Journal of muscle research and cell motility 9*, 359-366.

61) Katayama, E. (1989) The effects of various nucleotides on the structure of actin-

attached myosin subfragment-1 studied by quick-freeze deep-etch electron microscopy, *Journal of biochemistry 106*, 751-770.

62) Tokunaga M, Sutoh K, Wakabayashi T.(1991) Structure and structural change of the myosin head. *Adv Biophys, 27*, 157–167

63) J.W. Shriver, B. D. Sykes (1982) Energetics of the equilibrium between two nucleotide-free myosin subfragment 1 states using fluorine-19 Nuclear Magnetic Resonance, *Biochemistry 21,* 3022-3028.

64) Hiratsuka, T. (1992) Movement of Cys-697 in myosin ATPase associated withATP hydrolysis, *The Journal of biological chemistry 267*, 14941-14948.

65) Dalbey, R. E., Weiel, J., and Yount, R. G. (1983) Forster energy transfer measurements of thiol 1 to thiol 2 distances in myosin subfragment 1, *Biochemistry 22*, 4696-4706.

66) Wakabayashi, K., Tokunaga, M., Kohno, I., Sugimoto, Y., Hamanaka, T., Takezawa, Y., Wakabayashi, T., and Amemiya, Y. (1992) Small-angle synchrotron x-ray scattering reveals distinct shape changes of the myosin head during hydrolysis of
ATP, *Science (New York, N.Y 258*, 443-447

67) Dos Remedios, C. G., Millikan, R. G., and Morales, M. F. (1972) Polarization of tryptophan fluorescence from single striated muscle fibers. A molecular probe of contractile state, *The Journal of general physiology 59*, 103-120.

68) Borovikov, Y. S., and Levitsky, D. I. (1989) The effect of myosin light chain phosphorylation and Mg2+ on the conformation of myosin in thick filaments of glycerinated fibers of rabbit skeletal muscle, *European journal of biochemistry / FEBS 183*, 83-88.

69) Forkey, J. N., Quinlan, M. E., Shaw, M. A., Corrie, J. E., and Goldman, Y. E. (2003) Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization, *Nature 422*, 399-404.

70) Bobkova, E. A., Bobkov, A. A., Levitsky, D. I., and Reisler, E. (1999) Effects of SH1 and SH2 modifications on myosin: similarities and differences, *Biophysical journal 76*, 1001-1007.

71) Fisher, A. J., Smith, C. A., Thoden, J. B., Smith, R., Sutoh, K., Holden, H. M., and Rayment, I. (1995) X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4, *Biochemistry 34*, 8960-8972.

72) Gulick, A. M., Bauer, C. B., Thoden, J. B., and Rayment, I. (1997) X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes of the Dictyostelium

discoideum myosin motor domain, *Biochemistry 36*, 11619-11628.

73) Smith, C. A., and Rayment, I. (1996) X-ray structure of the magnesium(II).ADP.vanadate complex of the Dictyostelium discoideum myosin motor domain to 1.9 A resolution, *Biochemistry 35*, 5404-5417.

74) Godfrey, J. E., and Harrington, W. F. (1970) Self-association in the myosin system at high ionic strength. I. Sensitivity of the interaction to pH and ionic environment, *Biochemistry 9*, 886-893.

75) Spudich, J. A., and Watt, S. (1971) The regulation of rabbit skeletal muscle contraction. I. Biochemical studies of the interaction of the tropomyosin-troponin complex with actin and the proteolytic fragments of myosin, *The Journal of biological chemistry 246*, 4866-4871.

76) Weeds, A. G., and Pope, B. (1977) Studies on the chymotryptic digestion of myosin. Effects of divalent cations on proteolytic susceptibility, *Journal of molecular biology 111*, 129-157.

77) Ando, T., and Asai, H. (1980) Charge effects on the dynamic quenching of fluorescence of 1,N6-ethenoadenosine oligophosphates by iodide, thallium (I) and acrylamide, *Journal of biochemistry 88*, 255-264.

78) Chen, R. F., and Bowman, R. L. (1965) Fluorescence Polarization: Measurement with Ultraviolet-Polarizing Filters in a Spectrophotofluorometer, *Science (New York, N.Y 147*, 729-732.

79) Root, D. D., Shangguan, X., Xu, J., and McAllister, M. A. (1999) Determination of fluorescent probe orientations on biomolecules by conformational searching: algorithm testing and applications to the atomic model of myosin, *Journal of structural biology 127*, 22-34.

80) Joseph R. Lakowicz (1999) Principles of Fluorescence Spectroscopy , plenum press, New York.

81) Weber, G. (1952) Polarization of the fluorescence of macromolecules. I. Theory and experimental method, *The Biochemical journal 51*, 145-155.

82) Burghardt, T. P., and Ajtai, K. (1994) Following the rotational trajectory of the principal hydrodynamic frame of a protein using multiple probes, *Biochemistry 33*, 5376-5381.

83) Fajer, P. G. (1994) Determination of spin-label orientation within the myosin head, *Proceedings of the National Academy of Sciences of the United States of America 91*, 937-941.

84) E. Polak, (1970) Computational Methods in Optimisation, Academic Press,

New York .

85) McCammon, J. A., Gelin, B. R., and Karplus, M. (1977) Dynamics of folded proteins, *Nature 267*, 585-590.

86) Kay, L. E. (2005) NMR studies of protein structure and dynamics, *J Magn Reson 173*, 193-207.

87) Karplus, M., and Kuriyan, J. (2005) Molecular dynamics and protein function, *Proceedings of the National Academy of Sciences of the United States of America 102*, 6679-6685.

88) Smith, G. R., Sternberg, M. J., and Bates, P. A. (2005) The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking, *Journal of molecular biology 347*, 1077-1101.

89) Peters, G. H., Frimurer, T. M., Andersen, J. N., and Olsen, O. H. (1999) Molecular dynamics simulations of protein-tyrosine phosphatase 1B. I. ligand-induced changes in the protein motions, *Biophysical journal 77*, 505-515.

90) de Groot, B. L., van Aalten, D. M., Scheek, R. M., Amadei, A., Vriend, G., and Berendsen, H. J. (1997) Prediction of protein conformational freedom from distance constraints, *Proteins 29*, 240-251.

91) Amadei, A., Linssen, A. B., and Berendsen, H. J. (1993) Essential dynamics of proteins, *Proteins 17*, 412-425.

92) Schnell, J. R., Dyson, H. J., and Wright, P. E. (2004) Structure, dynamics, and catalytic function of dihydrofolate reductase, *Annual review of biophysics and biomolecular structure 33*, 119-140.

93) Fisher, A. J., Smith, C. A., Thoden, J., Smith, R., Sutoh, K., Holden, H. M., and Rayment, I. (1995) Structural studies of myosin:nucleotide complexes: a revised model for the molecular basis of muscle contraction, *Biophysical journal 68*, 19S-26S; discussion 27S-28S.

94) Botts, J., Thomason, J. F., and Morales, M. F. (1989) On the origin and transmission of force in actomyosin subfragment 1, *Proceedings of the National Academy of Sciences of the United States of America 86*, 2204-2208..

95) Sutoh, K., Tokunaga, M., and Wakabayashi, T. (1989) Electron microscopic microscopic mappings of myosin head with site-directed antibodies, *Journal of molecular biology 206*, 357-363.

96) Kull, F. J., Vale, R. D., and Fletterick, R. J. (1998) The case for a common ancestor: kinesin and myosin motor proteins and G proteins, *Journal of muscle research and cell motility 19*, 877-886.

97) Smith, C.A, and Rayment, I (1996) Active site comparisons highlight structural similarities between myosin and other P-loop proteins, *Biophysical journal 70,* 1590-1602.

98) Holmes, K. C., Angert, I., Kull, F. J., Jahn, W., and Schroder, R. R. (2003) Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide, *Nature 425*, 423-427.

99) Houdusse, A., Szent-Gyorgyi, A. G., and Cohen, C. (2000) Three conformational states of scallop myosin S1, *Proceedings of the National Academy of Sciences of the United States of America 97*, 11238-11243.

100) Reisler, E., Burke, M., and Harrington, W. F. (1974) Cooperative role of two sulfhydryl groups in myosin adenosine triphosphatase, *Biochemistry 13*, 20142022.

101) Wells, J. A., Knoeber, C., Sheldon, M. C., Werber, M. M., and Yount, R. G. (1980) Cross-linking of myosin subfragment 1. Nucleotide-enhanced modification by a variety of bifunctional reagents, *The Journal of biological chemistry 255*, 11135-11140.

102) Reisler, E., Burke, M., Himmelfarb, S., and Harrington, W. F. (1974) Spatial proximity of the two essential sulfhydryl groups of myosin, *Biochemistry 13*, 3837-3840.

103)Kinose, F., Wang, S. X., Kidambi, U. S., Moncman, C. L., and Winkelmann, D. A. (1996) Glycine 699 is pivotal for the motor activity of skeletal muscle myosin, *The Journal of cell biology 134*, 895-909.

104)Patterson, B., Ruppel, K. M., Wu, Y., and Spudich, J. A. (1997) Cold-sensitive mutants G680V and G691C of Dictyostelium myosin II confer dramatically different biochemical defects, *The Journal of biological chemistry 272*, 27612-27617.

105) Douglas. D, and Reisler, E (1992) Cooperativity of thiol-modified myosin filaments, ATPase and motility assays of myosin function, *Biophysical journal* 63, 730-740.

106) Walker, M., Zhang, X. Z., Jiang, W., Trinick, J., and White, H. D. (1999) Observation of transient disorder during myosin subfragment-1 binding to actin by stopped-flow fluorescence and millisecond time resolution electron cryomicroscopy: evidence that the start of the crossbridge power stroke in muscle has variable geometry, *Proceedings of the National Academy of Sciences of the*

United States of America 96, 465-470.

107) Roopnarine, O., Szent-Gyorgyi, A. G., and Thomas, D. D. (1998) Microsecond rotational dynamics of spin-labeled myosin regulatory light chain induced by relaxation and contraction of scallop muscle, *Biochemistry 37*, 1442814436.

108) Mulhern, S. A., and Eisenberg, E. (1978) Interaction of spin-labeled and N - (iodacetylaminoethyl)-5-naphthylamine-1-sulfonic acid SH1-blocked heavy meromyosin and myosin with actin and adenosine triphosphate, *Biochemistry 17*, 4419-4425.

109) Bobkov, A.A, Bobkova, E.A, Homsher, E, and Reisler, E. (1997) Activation of regulated actin by SH1-modified myosin subfragment 1, *Biochemistry 36,* 7733-7738.

110) Golitsina, N. L., Bobkov, A. A., Dedova, I. V., Pavlov, D. A., Nikolaeva, O. P., Orlov, V. N., and Levitsky, D. I. (1996) Differential scanning calorimetric study of the complexes of modified myosin subfragment 1 with ADP and vanadate or beryllium fluoride, *Journal of muscle research and cell motility 17*, 475-485.

111) Sasaki, N., Ohkura, R., and Sutoh, K. (2002) Dictyostelium myosin II as a model to study the actin-myosin interactions during force generation, *Journal of muscle research and cell motility 23*, 697-702.

112) Yengo, C. M., De La Cruz, E. M., Chrin, L. R., Gaffney, D. P., 2nd, and Berger, C. L. (2002) Actin-induced closure of the actin-binding cleft of smooth muscle myosin, *The Journal of biological chemistry 277*, 24114-24119.

113) Cook, R. K., Root, D., Miller, C., Reisler, E., and Rubenstein, P. A. (1993) Enhanced stimulation of myosin subfragment 1 ATPase activity by addition of negatively charged residues to the yeast actin NH2 terminus, *The Journal of biological chemistry 268*, 2410-2415.

114) Gawalapu, R., and Root, D. (2006) Fluorescence labeling and computational analysis of the strut of myosin's 50 kDa cleft, *Archives of Biochemistry and Biophysics 456,* 102-111

115) Text-book of Physical Chemistry By Samuel Glasstone Published 1940,D. Van Nostrand company, inc.

116) Field Desorption Mass Spectrometry (Practical Spectroscopy) By Laszlo Prokai.

117) Principles and Practice of Chromatography by RPW Scott..

118) Molecular modeling: Principles and applications by Andrew Leach.

119) Essentials Of Computational Chemistry: Theories And Models By Christopher J. Cramer

120) The Art of Molecular Dynamics Simulation By Dennis C. Rapaport